```
                        GenCore version 5.1.4_p5_4578
                    Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on:          March 22, 2003, 13:28:25 ; Search time 214 Seconds
                                            (without alignments)
                                            5094.597 Million cell updates/sec

Title:           US-09-784-358-2
Perfect score:   9346
Sequence:        1 MASWISPWWVLGMVFMHSP..........LNLQGLDRYKQKCSQSQET 1691

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters:      4569144

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Pending Patents AA_Main:*
                 1:  /cgn2_6/ptodata/1/paa/PCTUS_COMB.pep:*
                 2:  /cgn2_6/ptodata/1/paa/US06_COMB.pep:*
                 3:  /cgn2_6/ptodata/1/paa/US07_COMB.pep:*
                 4:  /cgn2_6/ptodata/1/paa/US080_COMB.pep:*
                 5:  /cgn2_6/ptodata/1/paa/US081_COMB.pep:*
                 6:  /cgn2_6/ptodata/1/paa/US082_COMB.pep:*
                 7:  /cgn2_6/ptodata/1/paa/US083_COMB.pep:*
                 8:  /cgn2_6/ptodata/1/paa/US084_COMB.pep:*
                 9:  /cgn2_6/ptodata/1/paa/US085_COMB.pep:*
                 10:  /cgn2_6/ptodata/1/paa/US086_COMB.pep:*
                 11:  /cgn2_6/ptodata/1/paa/US087_COMB.pep:*
                 12:  /cgn2_6/ptodata/1/paa/US088_COMB.pep:*
                 13:  /cgn2_6/ptodata/1/paa/US089_COMB.pep:*
                 14:  /cgn2_6/ptodata/1/paa/US090_COMB.pep:*
                 15:  /cgn2_6/ptodata/1/paa/US091_COMB.pep:*
                 16:  /cgn2_6/ptodata/1/paa/US092_COMB.pep:*
                 17:  /cgn2_6/ptodata/1/paa/US093_COMB.pep:*
                 18:  /cgn2_6/ptodata/1/paa/US094_COMB.pep:*
                 19:  /cgn2_6/ptodata/1/paa/US095_COMB.pep:*
                 20:  /cgn2_6/ptodata/1/paa/US096_COMB.pep:*
                 21:  /cgn2_6/ptodata/1/paa/US097_COMB.pep:*
                 22:  /cgn2_6/ptodata/1/paa/US098_COMB.pep:*
                 23:  /cgn2_6/ptodata/1/paa/US099_COMB.pep:*
                 24:  /cgn2_6/ptodata/1/paa/US100_COMB.pep:*
                 25:  /cgn2_6/ptodata/1/paa/US101_COMB.pep:*
                 26:  /cgn2_6/ptodata/1/paa/US102_COMB.pep:*
                 27:  /cgn2_6/ptodata/1/paa/US60_COMB.pep:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                                    SUMMARIES

                       %
Result               Query
  No.      Score     Match Length DB ID                          Description
-------------------------------------------------------------------------------

    1      9346      100.0   1691  21 US-09-784-358-2          Sequence 2, Appli
    2      9339       99.9   1691  21 US-09-789-390-4          Sequence 4, Appli
    3      8882       95.0   1617  21 US-09-784-358-16         Sequence 16, Appl
    4      7066       75.6   1281  21 US-09-789-390-5          Sequence 5, Appli
    5      7060       75.5   1281  21 US-09-789-390-2          Sequence 2, Appli
    6      5601       59.9   1023  21 US-09-789-390-61         Sequence 61, Appl
    7      5599       59.9   1023  21 US-09-789-390-25         Sequence 25, Appl
    8      5599       59.9   1023  21 US-09-789-390-27         Sequence 27, Appl
    9      5599       59.9   1023  21 US-09-789-390-28         Sequence 28, Appl
   10      5599       59.9   1023  21 US-09-789-390-59         Sequence 59, Appl
   11      5372       57.5   1036  27 US-60-245-228-264        Sequence 264, App
   12      5372       57.5   1036  27 US-60-258-272-108        Sequence 108, App
   13      5086       54.4    926  21 US-09-715-417A-28        Sequence 28, Appl
   14      4689       50.2    845  21 US-09-784-358-12         Sequence 12, Appl
   15      4225       45.2    771  21 US-09-784-358-14         Sequence 14, Appl
   16    3956.5       42.3   1762  24 US-10-044-807-2          Sequence 2, Appli
   17      3956       42.3    724   1 PCT-US02-10812-12       Sequence 12, Appl
   18      3956       42.3    724  21 US-09-784-358-8          Sequence 8, Appli
   19      3861       41.3   1745   1 PCT-US01-11988-2222     Sequence 2222, Ap
   20      3861       41.3   1745  22 US-09-800-723-89         Sequence 89, Appl
   21      3861       41.3   1745  22 US-09-833-245-2222       Sequence 2222, Ap
   22    3860.5       41.3   1766   1 PCT-US01-01349-554      Sequence 554, App
   23    3860.5       41.3   1766  21 US-09-764-853-554        Sequence 554, App
   24    3860.5       41.3   1766  24 US-10-072-326-554        Sequence 554, App
   25      3492       37.4    650  21 US-09-784-358-10         Sequence 10, Appl
   26      2891       30.9    874   1 PCT-US99-31025-66       Sequence 66, Appl
   27      2891       30.9    874  18 US-09-471-179-66         Sequence 66, Appl
   28      2891       30.9    874  21 US-09-796-753-20         Sequence 20, Appl
   29      2872       30.7    846   1 PCT-US99-31025-52       Sequence 52, Appl
   30      2872       30.7    846  18 US-09-471-179-52         Sequence 52, Appl
   31    2351.5       25.2    643  21 US-09-720-533-124        Sequence 124, App
   32    2351.5       25.2    643  27 US-60-112-129-12         Sequence 12, Appl
   33      2344       25.1    446  21 US-09-784-358-4          Sequence 4, Appli
   34      2058       22.0    525  17 US-09-380-138-301        Sequence 301, App
   35      2058       22.0    525  23 US-09-918-585A-301       Sequence 301, App
   36      2058       22.0    525  23 US-09-978-187B-301       Sequence 301, App
   37      2058       22.0    525  23 US-09-978-188A-301       Sequence 301, App
   38      2058       22.0    525  24 US-09-978-189-301        Sequence 301, App
   39      2058       22.0    525  23 US-09-978-190A-301       Sequence 301, App
   40      2058       22.0    525  23 US-09-978-191-301        Sequence 301, App
   41      2058       22.0    525  23 US-09-978-192A-301       Sequence 301, App
   42      2058       22.0    525  24 US-09-978-193A-301       Sequence 301, App
   43      2058       22.0    525  23 US-09-978-194A-301       Sequence 301, App
   44      2058       22.0    525  23 US-09-978-295A-301       Sequence 301, App
   45      2058       22.0    525  24 US-09-978-298A-301       Sequence 301, App

                                    ALIGNMENTS

RESULT 1
US-09-784-358-2
; Sequence 2, Application US/09784358
; GENERAL INFORMATION:
;  APPLICANT: Donoho, Gregory
;  APPLICANT: Scoville, John
;  APPLICANT: Turner, C. Alexander Jr.
;  APPLICANT: Friedrich, Glenn
;  APPLICANT: Zambrowicz, Brian
;  APPLICANT: Sands, Arthur T.
;  TITLE OF INVENTION: NOVEL HUMAN THROMBOSPONDIN REPEAT PROTEINS AND
;  TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING THE SAME
;  FILE REFERENCE: LEX-0134-USA
;  CURRENT APPLICATION NUMBER: US/09/784,358
;  CURRENT FILING DATE: 2001-02-15
;  PRIOR APPLICATION NUMBER: US 60/183,282
;  PRIOR FILING DATE: 2000-02-17
;  NUMBER OF SEQ ID NOS: 17
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 1691
;   TYPE: PRT
;   ORGANISM: homo sapiens
US-09-784-358-2

  Query Match             100.0%;  Score 9346;  DB 21;  Length 1691;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1691;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          1 MASWISPWWVLIGMVFMHSPLPQITAEKSPGAYFLPEFALSPQGSFLEDTIGEQFLTYRY 60
              ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db          1 MASWISPWWVLIGMVFMHSPLPQITAEKSPGAYFLPEFALSPQGSFLEDTIGEQFLTYRY 60

Qy         61 DDQISRNTRSDEDKDGNWDAWGDWSDCSRTCGGGASYSLRRCLTGRNCEGQNIRYKTCSN 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DDQISRNTRSDEDKDGNWDAWGDWSDCSRTCGGGASYSLRRCLTGRNCEGQNIRYKTCSN 120

Qy        121 HDCPPDAEDFRAQQCSAYNDVQYQGHYYEWLPRYNDPAAPCALKCHAQGQNLVVELAPKV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 HDCPPDAEDFRAQQCSAYNDVQYQGHYYEWLPRYNDPAAPCALKCHAQGQNLVVELAPKV 180

Qy        181 LDGIRCNIDSLDMCISGICQAVGCDRQLGSNAKEDNCGVCAGDGSTCRLVRGQSKSHVSP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 LDGIRCNIDSLDMCISGICQAVGCDRQLGSNAKEDNCGVCAGDGSTCRLVRGQSKSHVSP 240

Qy        241 EKREENVIAVPLGSRSVRITVKGPAHLFIESKTLQGSKGEHSFNSPGVFVVENTTVEFQR 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 EKREENVIAVPLGSRSVRITVKGPAHLFIESKTLQGSKGEHSFNSPGVFVVENTTVEFQR 300

Qy        301 GSERQTFKIPGPLMADFIFKTRYTAAKDSVVQFFFYQPISHQWRQTDFFPCTVTCGGGYQ 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GSERQTFKIPGPLMADFIFKTRYTAAKDSVVQFFFYQPISHQWRQTDFFPCTVTCGGGYQ 360

Qy        361 LNSAFCVDIRLKRVVPDHYCHYYPENVKPKPKLKECSMDPCPSSDGFKEIMPYDHFQPLP 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LNSAFCVDIRLKRVVPDHYCHYYPENVKPKPKLKECSMDPCPSSDGFKEIMPYDHFQPLP 420

Qy        421 RWEHNPWTACSVSCGGGIQRRSFVCVEESMHGEILQVEEWKCMYAPKPKVMQTCNLFDCP 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 RWEHNPWTACSVSCGGGIQRRSFVCVEESMHGEILQVEEWKCMYAPKPKVMQTCNLFDCP 480

Qy        481 KWIAMEWSQCTVTCGRGLRYRVVLCINHRGEHVGGCNPQLKLHIKEECVIPIPCYKPKEK 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 KWIAMEWSQCTVTCGRGLRYRVVLCINHRGEHVGGCNPQLKLHIKEECVIPIPCYKPKEK 540

Qy        541 SPVEAKLPWLKQAQELEETRIATEEPTFIPEPWSACSTTCGPGVQVREVKCRVLLTFTQT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 SPVEAKLPWLKQAQELEETRIATEEPTFIPEPWSACSTTCGPGVQVREVKCRVLLTFTQT 600

Qy        601 ETELPEEECEGPKLPTERPCLLEACDESPASRELDIPLPEDSETTYDWEYAGFTPCTATC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ETELPEEECEGPKLPTERPCLLEACDESPASRELDIPLPEDSETTYDWEYAGFTPCTATC 660

Qy        661 LGGHQEAIAVCLHIQTQQTVNDSLCDMVHRPPAMSQACNTEPCPPRWHVGSWGPCSATCG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 LGGHQEAIAVCLHIQTQQTVNDSLCDMVHRPPAMSQACNTEPCPPRWHVGSWGPCSATCG 720

Qy        721 VGIQTRDVYCLHPGETPAPPEECRDEKPHALQACNQFDCPPGWHIEEWQQCSRTCGGGTQ 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 VGIQTRDVYCLHPGETPAPPEECRDEKPHALQACNQFDCPPGWHIEEWQQCSRTCGGGTQ 780

Qy        781 NRRVTCRQLLTDGSFLNLSDELCQGPKASSHKSCARTDCPPHLAVGDWSKCSVSCGVGIQ 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 NRRVTCRQLLTDGSFLNLSDELCQGPKASSHKSCARTDCPPHLAVGDWSKCSVSCGVGIQ 840

Qy        841 RRKQVCQRLAAKGRRIPLSEMMCRDLPCFPLVRSCQMPECSKIKSEMKTKLGEQGPQILS 900
              |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db        841 RRKQVCQRLAAKGRRIPLSEMMCRDLPGFPLVRSCQMPECSKIKSEMKTKLGEQGPQILS 900

Qy        901 VQRVYIQTRFFFKRINLTIGSRAYLLPNTSVIIKCPVRRFQKSLIQWEKDGRCLQNSKRLG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 VQRVYIQTRFFFKRINLTIGSRAYLLPNTSVIIKCPVRRFQKSLIQWEKDGRCLQNSKRLG 960

Qy        961 IIKSGSLKIHGLAAPDIGVYRCIAGSAQETVVLKLIGTDNRLIARPALREPMREYPGMDH 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 IIKSGSLKIHGLAAPDIGVYRCIAGSAQETVVLKLIGTDNRLIARPALREPMREYPGMDH 1020

Qy       1021 SEANSLGVTWHKMRQMWNNKNDLYLDDDHISNQPFLRALLGHCSNSAGSTNSWELKNKQF 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 SEANSLGVTWHKMRQMWNNKNDLYLDDDHISNQPFLRALLGHCSNSAGSTNSWELKNKQF 1080

Qy       1081 EAAVKQGAYSMDTAQFDELIRNMSQLMETGEVSDDLASQLIYQLVAELAKAQPTHMQWRG 1140
```

```
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 EAAVKQGAYSMDTAQFDELIRNMSQLMETGEVSDDLASQLIYQLVAELAKAQPTHMQWRG 1140

Qy       1141 IQEETPPAAQLRGETGSVSQSSHAKNSGKLTFKPKGPVLMRQSQPPSISFNKTINSRIGN 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 IQEETPPAAQLRGETGSVSQSSHAKNSGKLTFKPKGPVLMRQSQPPSISFNKTINSRIGN 1200

Qy       1201 TVYITKRTEVINILCDLITPSEATYTWTKDGTLLQPSVKIILDGTGKIQIQNPTRKEQGI 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TVYITKRTEVINILCDLITPSEATYTWTKDGTLLQPSVKIILDGTGKIQIQNPTRKEQGI 1260

Qy       1261 YECSVANHLGSDVESSSVLYAEAPVILSVERNITKPEHNHLSVVVGGIVEAALGANVTIR 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 YECSVANHLGSDVESSSVLYAEAPVILSVERNITKPEHNHLSVVVGGIVEAALGANVTIR 1320

Qy       1321 CPVKGVPQPNITWLKRGGSLSGNVSLLFNGSLLLQNVSLENEGTYVCIATNALGKAVATS 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CPVKGVPQPNITWLKRGGSLSGNVSLLFNGSLLLQNVSLENEGTYVCIATNALGKAVATS 1380

Qy       1381 VLHLLERRWPESRIVFLQGHKKYILQATNTRTNSNDPTGEPPPQEPFWEPGNWSHCSATC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 VLHLLERRWPESRIVFLQGHKKYILQATNTRTNSNDPTGEPPPQEPFWEPGNWSHCSATC 1440

Qy       1441 GHLGARIQRPQCVMANGQEVSEALCDHLQKPLAGFEPCNIRDCPARWFTSVWSQCSVSCG 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GHLGARIQRPQCVMANGQEVSEALCDHLQKPLAGFEPCNIRDCPARWFTSVWSQCSVSCG 1500

Qy       1501 EGYHSRQVTCKRTKANGTVQVVSPRACAPKDRPLGRKPCFGHPCVQWEPGNRCPGRCMGR 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 EGYHSRQVTCKRTKANGTVQVVSPRACAPKDRPLGRKPCFGHPCVQWEPGNRCPGRCMGR 1560

Qy       1561 AVRMQQRHTACQHNSSDSNCDDRKRPTLRRNCTSGACDVCWHTGPWKPCTAACGRGFQSR 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 AVRMQQRHTACQHNSSDSNCDDRKRPTLRRNCTSGACDVCWHTGPWKPCTAACGRGFQSR 1620

Qy       1621 KVDCIHTRSCKPVAKRHCVQKKKPISWRHCLGPSCDRDCTDTTHYCMFVKHLNLCSLDRY 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 KVDCIHTRSCKPVAKRHCVQKKKPISWRHCLGPSCDRDCTDTTHYCMFVKHLNLCSLDRY 1680

Qy       1681 KQRCCQSCQEG 1691
              |||||||||||
Db       1681 KQRCCQSCQEG 1691
```

```
RESULT 2
US-09-789-390-4
; Sequence 4, Application US/09789390
; GENERAL INFORMATION:
;  APPLICANT: Vernet, Corine
;  APPLICANT:  Fernandes, Elma
;  APPLICANT:  MacDougall, John
;  APPLICANT:  Shimkets, Richard A
;  APPLICANT:  Spaderna, Steven K
;  TITLE OF INVENTION: NOVEL POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
;  FILE REFERENCE: 15966-692
;  CURRENT APPLICATION NUMBER: US/09/789,390
;  CURRENT FILING DATE:  2001-02-23
;  PRIOR APPLICATION NUMBER: 60/185,548
;  PRIOR FILING DATE: 2000-02-28
;  PRIOR APPLICATION NUMBER: 60/199,957
;  PRIOR FILING DATE: 2000-04-27
;  PRIOR APPLICATION NUMBER: 60/184,951
;  PRIOR FILING DATE: 2000-02-25
;  PRIOR APPLICATION NUMBER: 60/185,967
;  PRIOR FILING DATE: 2000-03-01
;  PRIOR APPLICATION NUMBER: 60/197,723
;  PRIOR FILING DATE: 2000-04-18
;  NUMBER OF SEQ ID NOS: 77
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
;   LENGTH: 1691
;   TYPE: PRT
;   ORGANISM: Homo sapiens
```

```
US-09-789-390-4

  Query Match             99.9%;  Score 9339;  DB 21;  Length 1691;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1689;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MASWISPWWVLIGMVFMHSPLPQTTAEKSPGAYFLPEFALSPQGSFLEDTTGEQFLIYRY 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MASWISPWWVLIGMVFMHSPLPQTTAEKSPGAYFLPEFALSPQGSFLEDTTGEQFLIYRY 60

Qy         61 DDQTSRNTRSDEDKDGNWDAWGDWSDCSRTCGGGASYSLRRCLTGRNCEGQNIRYKTCSN 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DDQTSRNTRSDEDKDGNWDAWGDWSDCSRTCGGGASYSLRRCLTGRNCEGQNIRYKTCSN 120

Qy        121 HDCPPDAEDFRAQQCSAYNDVQYQGHYYEWLPRYNDPAAPCALKCHAQGQNLVVELAPKV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 HDCPPDAEDFRAQQCSAYNDVQYQGHYYEWLPRYNDPAAPCALKCHAQGQNLVVELAPKV 180

Qy        181 LDGTRCNTDSLDMCISGICQAVGCDRQLGSNAKEDNCGVCAGDGSTCRLVRGQSKSHVSP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 LDGTRCNTDSLDMCISGICQAVGCDRQLGSNAKEDNCGVCAGDGSTCRLVRGQSKSHVSP 240

Qy        241 EKREENVIAVPLGSRSVRITVKGPAHLFIESKTLQGSKGEHSFNSPGVFVVENTTVEFQR 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 EKREENVIAVPLGSRSVRITVKGPAHLFIESKTLQGSKGEHSFNSPGVFVVENTTVEFQR 300

Qy        301 GSERQTFKIPGPLMADFIFKTRYTAAKDSVVQFFFYQPISHQWRQTDFFPCTVTCGGGYQ 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GSERQTFKIPGPLMADFIFKTRYTAAKDSVVQFFFYQPISHQWRQTDFFPCTVTCGGGYQ 360

Qy        361 LNSAECVDIRLKRVVPDHYCHYYPENVKPKPKLKECSMDPCPSSDGFKEIMPYDHFQPLP 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LNSAECVDIRLKRVVPDHYCHYYPENVKPKPKLKECSMDPCPSSDGFKEIMPYDHFQPLP 420

Qy        421 RWEHNPWTACSVSCGGGIQRRSFVCVEESMHGEILQVEEWKCMYAPKPKVMQTCNLFDCP 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 RWEHNPWTACSVSCGGGIQRRSFVCVEESMHGEILQVEEWKCMYAPKPKVMQTCNLFDCP 480

Qy        481 KWIAMEWSQCTVTCGRGLRYRVVLCINHRGEHVGGCNPQLKLHIKEECVIPIPCYKPKEK 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 KWIAMEWSQCTVTCGRGLRYRVVLCINHRGEHVGGCNPQLKLHIKEECVIPIPCYKPKEK 540

Qy        541 SPVEAKLPWLKQAQELEETRIATEEPTFIPEPWSACSTTCGPGVQVREVKCRVLLTFTQT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 SPVEAKLPWLKQAQELEETRIATEEPTFIPEPWSACSTTCGPGVQVREVKCRVLLTFTQT 600

Qy        601 ETELPEEECEGPKLPTERPCLLEACDESPASRELDIPLPEDSETTYDWEYAGFTPCTATC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ETELPEEECEGPKLPTERPCLLEACDESPASRELDIPLPEDSETTYDWEYAGFTPCTATC 660

Qy        661 IGGHQEATAVCLHIQTQQTVNDSLCDMVHRPPAMSQACNTEPCPPRWHVGSWGPCSATCG 720
              :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 VGGHQEATAVCLHIQTQQTVNDSLCDMVHRPPAMSQACNTEPCPPRWHVGSWGPCSATCG 720

Qy        721 VGTQTRDVYCLHPGETPAPPEECRDEKPHALQACNQFDCPPGWHIEEWQQCSRTCGGGTQ 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 VGTQTRDVYCLHPGETPAPPEECRDEKPHALQACNQFDCPPGWHIEEWQQCSRTCGGGTQ 780

Qy        781 NRRVTCRQLLTDGSFLNLSDELCQGPKASSHKSCARTDCPPHLAVGDWSKCSVSCGVGIQ 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 NRRVTCRQLLTDGSFLNLSDELCQGPKASSHKSCARTDCPPHLAVGDWSKCSVSCGVGIQ 840

Qy        841 RRKQVCQRLAAKGRRIPLSEMMCRDLPGFPLVRSCQMPECSKIKSEMKTKLGEQGPQILS 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 RRKQVCQRLAAKGRRIPLSEMMCRDLPGFPLVRSCQMPECSKIKSEMKTKLGEQGPQILS 900

Qy        901 VQRVYIQTREEKRINLTIGSRAYLLPNTSVIIKCPVRRFQKSLIQWEKDGRCLQNSKRLG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 VQRVYIQTREEKRINLTIGSRAYLLPNTSVIIKCPVRRFQKSLIQWEKDGRCLQNSKRLG 960

Qy        961 ITKSGSLKIHGLAAPDIGVYRCIAGSAQETVVLKLIGTDNRLIARPALREPMREYPGMDH 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 ITKSGSLKIHGLAAPDIGVYRCIAGSAQETVVLKLIGTDNRLIARPALREPMREYPGMDH 1020

Qy       1021 SEANSLGVTWHKMRQMWNNKNDLYLDDDHISNQPFLRALLGHCSNSAGSTNSWELKNKQF 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 SEANSLGVTWHKMRQMWNNKNDLYLDDDHISNQPFLRALLGHCSNSAGSTNSWELKNKQF 1080

Qy       1081 EAAVKQGAYSMDTAQFDELIRNMSQLMETGEVSDDLASQLIYQLVAELAKAQPTHMQWRG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 EAAVKQGAYSMDTAQFDELIRNMSQLMETGEVSDDLASQLIYQLVAELAKAQPTHMQWRG 1140

Qy       1141 IQEETPPAAQLRGETGSVSQSSHAKNSGKLTFKPKGPVLMRQSQPPSISFNKTINSRIGN 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 IQEETPPAAQLRGETGSVSQSSHAKNSGKLTFKPKGPVLMRQSQPPSISFNKTINSRIGN 1200

Qy       1201 TVYITKRTEVINILCDLITPSEATYTWTKDGTLLQPSVKIILDGTGKIQIQNPTRKEQGI 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TVYITKRTEVINILCDLITPSEATYTWTKDGTLLQPSVKIILDGTGKIQIQNPTRKEQGI 1260

Qy       1261 YECSVANHLGSDVESSSVLYAEAPVILSVERNITKPEHNHLSVVVGGIVEAALGANVTIR 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 YECSVANHLGSDVESSSVLYAEAPVILSVERNITKPEHNHLSVVVGGIVEAALGANVTIR 1320

Qy       1321 CPVKGVPQPNITWLKRGGSLSGNVSLLFNGSLLLQNVSLENEGTYVCIATNALGKAVATS 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CPVKGVPQPNITWLKRGGSLSGNVSLLFNGSLLLQNVSLENEGTYVCIATNALGKAVATS 1380

Qy       1381 VLHLLERRWPESRIVFLQGHKKYILQATNTRTNSNDPTGEPPPQEPFWEPGNWSHCSATC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 VLHLLERRWPESRIVFLQGHKKYILQATNTRTNSNDPTGEPPPQEPFWEPGNWSHCSATC 1440

Qy       1441 GHLGARIQRPQCVMANGQEVSEALCDHLQKPLAGFEPCNIRDCPARWFTSVWSQCSVSCG 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GHLGARIQRPQCVMANGQEVSEALCDHLQKPLAGFEPCNIRDCPARWFTSVWSQCSVSCG 1500

Qy       1501 EGYHSRQVTCKRTKANGTVQVVSPRACAPKDRPLGRKPCFGHPCVQWEPGNRCPGRCMGR 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 EGYHSRQVTCKRTKANGTVQVVSPRACAPKDRPLGRKPCFGHPCVQWEPGNRCPGRCMGR 1560

Qy       1561 AVRMQQRHTACQHNSSDSNCDDRKRPTLRRNCTSGACDVCWHTGPWKPCTAACGRGFQSR 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 AVRMQQRHTACQHNSSDSNCDDRKRPTLRRNCTSGACDVCWHTGPWKPCTAACGRGFQSR 1620

Qy       1621 KVDCIHTRSCKPVAKRHCVQKKKPISWRHCLGPSCDRDCTDTTHYCMFVKHLNLCSLDRY 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 KVDCIHTRSCKPVAKRHCVQKKKPISWRHCLGPSCDRDCTDTTHYCMFVKHLNLCSLDRY 1680

Qy       1681 KQRCCQSCQEG 1691
              |||||||||||
Db       1681 KQRCCQSCQEG 1691


RESULT 3
US-09-784-358-16
; Sequence 16, Application US/09784358
; GENERAL INFORMATION:
;  APPLICANT: Donoho, Gregory
;  APPLICANT:  Scoville, John
;  APPLICANT:  Turner, C. Alexander Jr.
;  APPLICANT:  Friedrich, Glenn
;  APPLICANT:  Zambrowicz, Brian
;  APPLICANT:  Sands, Arthur T.
;  TITLE OF INVENTION: NOVEL HUMAN THROMBOSPONDIN REPEAT PROTEINS AND
;  TITLE OF INVENTION:  POLYNUCLEOTIDES ENCODING THE SAME
;  FILE REFERENCE: LEX-0134-USA
;  CURRENT APPLICATION NUMBER: US/09/784,358
;  CURRENT FILING DATE:  2001-02-15
;  PRIOR APPLICATION NUMBER: US 60/183,282
;  PRIOR FILING DATE: 2000-02-17
;  NUMBER OF SEQ ID NOS: 17
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16
;   LENGTH: 1617
;   TYPE: PRT
;   ORGANISM: homo sapiens
```

```
US-09-784-358-16

  Query Match             95.0%;  Score 8882;  DB 21,  Length 1617;
  Best Local Similarity   95.6%;  Pred. No. 0;
  Matches 1617;  Conservative    0;  Mismatches    0;  Indels   74;  Gaps    1;

Qy          1 MASWTSPWWVLIGMVFMHSPLPQTTAEKSPGAYFLPEFALSPQGSFLEDTTGEQFLTYRY 60
              ||||||||||||||||||||||||||||||
Db          1 MASWTSPWWVLIGMVFMHSPLPQTTAEKSP------------------------------ 30

Qy         61 DDQTSRNIRSDEDKDGNWDAWGDWSDCSRTCGGGASYSLRRCLTGRNCEGQNIRYKTCSN 120
                                                          ||||||||||||||||
Db         31 --------------------------------------------GRNCEGQNIRYKTCSN 46

Qy        121 HDCFPDAEDFRAQQCSAYNDVQYQGHYYEWLPRYNDPAAPCALKCHAQGQNLVVELAPKV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         47 HDCFPDAEDFRAQQCSAYNDVQYQGHYYEWLPRYNDPAAPCALKCHAQGQNLVVELAPKV 106

Qy        181 LDGTRCNTDSLDMCISGICQAVGCDRQLGSNAKEDNCGVCAGDGSTCRLVRGQSKSHVSP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 LDGTRCNTDSLDMCISGICQAVGCDRQLGSNAKEDNCGVCAGDGSTCRLVRGQSKSHVSP 166

Qy        241 EKREENVIAVPLGSRSVRITVKGPAHLFIESKTLQGSKGEHSFNSPGVFVVENTTVEFQR 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 EKREENVIAVPLGSRSVRITVKGPAHLFIESKTLQGSKGEHSFNSPGVFVVENTTVEFQR 226

Qy        301 GSERQTFKIPGPLMADFIFKTRYTAAKDSVVQFFFYQPISHQWRQTDFFPCTVTCGGGYQ 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 GSERQTFKIPGPLMADFIFKTRYTAAKDSVVQFFFYQPISHQWRQTDFFPCTVTCGGGYQ 286

Qy        361 LNSAECVDIRLKRVVPDHYCHYYPENVKPKPKLKECSMDPCPSSDGFKEIMPYDHFQPLP 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 LNSAECVDIRLKRVVPDHYCHYYPENVKPKPKLKECSMDPCPSSDGFKEIMPYDHFQPLP 346

Qy        421 RWEHNPWTACSVSCGGGIQRRSFVCVEESMHGEILQVEEWKCMYAPKPKVMQTCNLFDCP 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        347 RWEHNPWTACSVSCGGGIQRRSFVCVEESMHGEILQVEEWKCMYAPKPKVMQTCNLFDCP 406

Qy        481 KWIAMEWSQCTVTCGRGLRYRVVLCINHRGEHVGGCNPQLKLHIKEECVIPIPCYKPKEK 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        407 KWIAMEWSQCTVTCGRGLRYRVVLCINHRGEHVGGCNPQLKLHIKEECVIPIPCYKPKEK 466

Qy        541 SPVEAKLPWLKQAQELEETRIATEEPTFIPEPWSACSTTCGPGVQVREVKCRVLLTFTQT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        467 SPVEAKLPWLKQAQELEETRIATEEPTFIPEPWSACSTTCGPGVQVREVKCRVLLTFTQT 526

Qy        601 ETELPEEECEGPKLPTERPCLLEACDESPASRELDIPLPEDSETTYDWEYAGFTPCTATC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        527 ETELPEEECEGPKLPTERPCLLEACDESPASRELDIPLPEDSETTYDWEYAGFTPCTATC 586

Qy        661 LGGHQEAIAVCLHIQTQQTVNDSLCDMVHRPPAMSQACNTEPCPPRWHVGSWGPCSATCG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        587 LGGHQEAIAVCLHIQTQQTVNDSLCDMVHRPPAMSQACNTEPCPPRWHVGSWGPCSATCG 646

Qy        721 VGIQTRDVYCLHPGETPAPPEECRDEKPHALQACNQFDCPPGWHIEEWQQCSRTCGGGTQ 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        647 VGIQTRDVYCLHPGETPAPPEECRDEKPHALQACNQFDCPPGWHIEEWQQCSRTCGGGTQ 706

Qy        781 NRRVTCRQLLTDGSFLNLSDELCQGPKASSHKSCARTDCPPHLAVGDWSKCSVSCGVGIQ 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        707 NRRVTCRQLLTDGSFLNLSDELCQGPKASSHKSCARTDCPPHLAVGDWSKCSVSCGVGIQ 766

Qy        841 RREQVCQRLAAKGRRIPLSEMMCRDLPGFPLVRSCQMPECSKIKSEMKTKLGEQGPQILS 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        767 RREQVCQRLAAKGRRIPLSEMMCRDLPGFPLVRSCQMPECSKIKSEMKTKLGEQGPQILS 826

Qy        901 VQRVYIQTREEKRINLTIGSRAYLLPNTSVIIKCPVRRFQKSLIQWEKDGRCLQNSKRLG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        827 VQRVYIQTREEKRINLTIGSRAYLLPNTSVIIKCPVRRFQKSLIQWEKDGRCLQNSKRLG 886

Qy        961 ITKSGSLKIHGLAAPDIGVYRCIAGSAQETVVLKLIGTDNRLIARPALREPMREYPGMDH 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        887 ITKSGSLKIHGLAAPDIGVYRCIAGSAQETVVLKLIGTDNRLIARPALREPMREYPGMDH 946

Qy       1021 SEANSLGVTWHKMRQMWNNKNDLYLDDDHISNQPFLRALLGHCSNSAGSTNSWELKNKQF 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        947 SEANSLGVTWHKMRQMWNNKNDLYLDDDHISNQPFLRALLGHCSNSAGSTNSWELKNKQF 1006

Qy       1081 EAAVKQGAYSMDTAQFDELIRNMSQIMETGEVSDDLASQLIYQLVAELAKAQPTHMQWRG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1007 EAAVKQGAYSMDTAQFDELIRNMSQIMETGEVSDDLASQLIYQLVAELAKAQPTHMQWRG 1066

Qy       1141 IQEETPPAAQLRGETGSVSQSSHAKNSGKLTFKPKGPVLMRQSQPPSISFNKTINSRIGN 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1067 IQEETPPAAQLRGETGSVSQSSHAKNSGKLTFKPKGPVLMRQSQPPSISFNKTINSRIGN 1126

Qy       1201 TVYITKRTEVINILCDLITPSEATYTWTKDGTLLQPSVKIIIDGTGKIQIQNPTRKEQGI 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1127 TVYITKRTEVINILCDLITPSEATYTWTKDGTLLQPSVKIIIDGTGKIQIQNPTRKEQGI 1186

Qy       1261 YECSVANHLGSDVESSSVLYAEAPVILSVERNITKPEHNHLSVVVGGIVEAALGANVTIR 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1187 YECSVANHLGSDVESSSVLYAEAPVILSVERNITKPEHNHLSVVVGGIVEAALGANVTIR 1246

Qy       1321 CPVKGVPQPNITWLKRGGSLSGNVSLLFNGSLLLQNVSLENEGTYVCIATNALGKAVATS 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1247 CPVKGVPQPNITWLKRGGSLSGNVSLLFNGSLLLQNVSLENEGTYVCIATNALGKAVATS 1306

Qy       1381 VLHLLERRWPESRIVFLQGHKKYILQATNTRTNSNDPTGEPPPQEPFWEPGNWSHCSATC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1307 VLHLLERRWPESRIVFLQGHKKYILQATNTRTNSNDPTGEPPPQEPFWEPGNWSHCSATC 1366

Qy       1441 GHLGARIQRPQCVMANGQEVSEALCDHLQKPLAGFEPCNIRDCPARWFTSVWSQCSVSCG 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1367 GHLGARIQRPQCVMANGQEVSEALCDHLQKPLAGFEPCNIRDCPARWFTSVWSQCSVSCG 1426

Qy       1501 EGYHSRQVTCKRTKANGTVQVVSPRACAPKDRPLGRKPCFGHPCVQWEPGNRCPGRCMGR 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1427 EGYHSRQVTCKRTKANGTVQVVSPRACAPKDRPLGRKPCFGHPCVQWEPGNRCPGRCMGR 1486

Qy       1561 AVRMQQRHTACQHNSSDSNCDDRKRPTLRRNCTSGACDVCWHTGPWKPCTAACGRGFQSR 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1487 AVRMQQRHTACQHNSSDSNCDDRKRPTLRRNCTSGACDVCWHTGPWKPCTAACGRGFQSR 1546

Qy       1621 KVDCIHTRSCKPVAKRHCVQKKKPISWRHCLGPSCDRDCTDTTHYCMFVKHLNLCSLDRY 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1547 KVDCIHTRSCKPVAKRHCVQKKKPISWRHCLGPSCDRDCTDTTHYCMFVKHLNLCSLDRY 1606

Qy       1681 KQRCCQSCQEG 1691
              |||||||||||
Db       1607 KQRCCQSCQEG 1617


RESULT 4
US-09-789-390-5
; Sequence 5, Application US/09789390
; GENERAL INFORMATION:
;  APPLICANT: Vernet, Corine
;  APPLICANT:  Fernandes, Elma
;  APPLICANT:  MacDougall, John
;  APPLICANT:  Shimkets, Richard A
;  APPLICANT:  Spaderna, Steven K
;  TITLE OF INVENTION: NOVEL POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
;  FILE REFERENCE: 15966-692
;  CURRENT APPLICATION NUMBER: US/09/789,390
;  CURRENT FILING DATE:  2001-02-23
;  PRIOR APPLICATION NUMBER: 60/185,548
;  PRIOR FILING DATE: 2000-02-28
;  PRIOR APPLICATION NUMBER: 60/199,957
;  PRIOR FILING DATE: 2000-04-27
;  PRIOR APPLICATION NUMBER: 60/184,951
;  PRIOR FILING DATE: 2000-02-25
;  PRIOR APPLICATION NUMBER: 60/185,967
;  PRIOR FILING DATE: 2000-03-01
;  PRIOR APPLICATION NUMBER: 60/197,723
;  PRIOR FILING DATE: 2000-04-18
;  NUMBER OF SEQ ID NOS: 77
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
;   LENGTH: 1281
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-789-390-5

  Query Match             75.6%;  Score 7066;  DB 21;  Length 1281;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1280;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        411 MPYDHFQPLPRWEHNPWTACSVSCGGGIQRRSFVCVEESMHGEILQVEEWKCMYAPKPKV 470
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MPYDHFQPLPRWEHNPWTACSVSCGGGIQRRSFVCVEESMHGEILQVEEWKCMYAPKPKV 60

Qy        471 MQTCNLFDCPKWIAMEWSQCTVTCGRGLRYRVVLCINHRGEHVGGCNPQLKLHIKEECVI 530
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 MQTCNLFDCPKWIAMEWSQCTVTCGRGLRYRVVLCINHRGEHVGGCNPQLKLHIKEECVI 120

Qy        531 PIPCYKPKEKSPVEAKLPWLKQAQELEETRIATEEPTFIPEPWSACSTTCGPGVQVREVK 590
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 PIPCYKPKEKSPVEAKLPWLKQAQELEETRIATEEPTFIPEPWSACSTTCGPGVQVREVK 180

Qy        591 CRVLLTFTQTETELPEEECEGPKLPTERPCLLEACDESPASRELDIPLPEDSETTYDWEY 650
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CRVLLTFTQTETELPEEECEGPKLPTERPCLLEACDESPASRELDIPLPEDSETTYDWEY 240

Qy        651 AGFTPCTATCVGGHQEATAVCLHIQTQQTVNDSLCDMVHRPPAMSQACNTEPCPPRWHVG 710
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGFTPCTATCVGGHQEATAVCLHIQTQQTVNDSLCDMVHRPPAMSQACNTEPCPPRWHVG 300

Qy        711 SWGPCSATCGVGIQTRDVYCLHPGETPAPPEECRDEKPHALQACNQFDCPPGWHIEEWQQ 770
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 SWGPCSATCGVGIQTRDVYCLHPGETPAPPEECRDEKPHALQACNQFDCPPGWHIEEWQQ 360

Qy        771 CSRTCGGGTQNRRVTCRQLLTDGSFLNLSDELCQGPKASSHKSCARTDCPPHLAVGDWSK 830
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CSRTCGGGTQNRRVTCRQLLTDGSFLNLSDELCQGPKASSHKSCARTDCPPHLAVGDWSK 420

Qy        831 CSVSCGVGIQRRKQVCQRLAAKGRRIPLSEMMCRDLPGFPLVRSCQMPECSKIKSEMKTK 890
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CSVSCGVGIQRRKQVCQRLAAKGRRIPLSEMMCRDLPGFPLVRSCQMPECSKIKSEMKTK 480

Qy        891 LGEQGPQILSVQRVYIQTREEKRINLTIGSRAYLLPNTSVIIKCPVRRFQKSLIQWEKDG 950
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 LGEQGPQILSVQRVYIQTREEKRINLTIGSRAYLLPNTSVIIKCPVRRFQKSLIQWEKDG 540

Qy        951 RCLQNSKRLGITKSGSLKIHGLAAPDIGVYRCIAGSAQETVVLKLIGTDNRLIARPALRE 1010
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:
Db        541 RCLQNSKRLGITKSGSLKIHGLAAPDIGVYRCIAGSAQETVVLKLIGTDNRLIARPALRP 600

Qy       1011 PMREYPGMDHSEANSLGVTWHKMRQMWNNKNDLYLDDDHISNQPFLRALLGHCSNSAGST 1070
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 PMREYPGMDHSEANSLGVTWHKMRQMWNNKNDLYLDDDHISNQPFLRALLGHCSNSAGST 660

Qy       1071 NSWELKNKQFEAAVKQGAYSMDTAQFDELIRNMSQLMETGEVSDDLASQLIYQLVAELAK 1130
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 NSWELKNKQFEAAVKQGAYSMDTAQFDELIRNMSQLMETGEVSDDLASQLIYQLVAELAK 720

Qy       1131 AQPTHMQWRGIQEEIPPAAQLRGETGSVSQSSHAKNSGKLTFKPKGPVLMRQSQPPSISF 1190
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AQPTHMQWRGIQEEIPPAAQLRGETGSVSQSSHAKNSGKLTFKPKGPVLMRQSQPPSISF 780

Qy       1191 NKTINSRIGNTVYITKRTEVINILCDLITPSEATYTWTKDGTLLQPSVKIILDGTGKIQI 1250
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 NKTINSRIGNTVYITKRTEVINILCDLITPSEATYTWTKDGTLLQPSVKIILDGTGKIQI 840

Qy       1251 QNPTRKEQGIYECSVANHLGSDVESSSVLYAEAPVILSVERNITKPEHNHLSVVVGGIVE 1310
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 QNPTRKEQGIYECSVANHLGSDVESSSVLYAEAPVILSVERNITKPEHNHLSVVVGGIVE 900

Qy       1311 AALGANVTIRCPVKGVPQPNITWLKRGGSLSGNVSLLFNGSLLLQNVSLENEGTYVCIAT 1370
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AALGANVTIRCPVKGVPQPNITWLKRGGSLSGNVSLLFNGSLLLQNVSLENEGTYVCIAT 960

Qy       1371 NALGKAVATSVLHLLERRWPESRIVFLQGHKKYILQATNTRTNSNDPTGEPPPQEPFWEP 1430
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 NALGKAVATSVLHLLERRWPESRIVFLQGHKKYILQATNTRTNSNDPTGEPPPQEPFWEP 1020

Qy       1431 GNWSHCSATCGHLGARIQRPQCVMANGQEVSEALCDHLQKPLAGFEPCNIRDCPARWFTS 1490
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GNWSHCSATCGHLGARIQRPQCVMANGQEVSEALCDHLQKPLAGFEPCNIRDCPARWFTS 1080

Qy       1491 VWSQCSVSCGEGYHSRQVTCKRTKANGTVQVVSPRACAPKDRPLGRKPCFGHPCVQWEPG 1550
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 VWSQCSVSCGEGYHSRQVTCKRTKANGTVQVVSPRACAPKDRPLGRKPCFGHPCVQWEPG 1140

Qy       1551 NRCPGRCMGRAVRMQQRHTACQHNSSDSNCDDRKRPTLRRNCTSGACDVCWHTGPWKPCT 1610
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 NRCPGRCMGRAVRMQQRHTACQHNSSDSNCDDRKRPTLRRNCTSGACDVCWHTGPWKPCT 1200

Qy       1611 AACGRGFQSRKVDCIHTRSCKPVAKRHCVQKKKPISWRHCLGPSCDRDCTDTTHYCMFVK 1670
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 AACGRGFQSRKVDCIHTRSCKPVAKRHCVQKKKPISWRHCLGPSCDRDCTDTTHYCMFVK 1260

Qy       1671 HLNLCSLDRYKQRCCQSCQEG 1691
              |||||||||||||||||||||
Db       1261 HLNLCSLDRYKQRCCQSCQEG 1281


RESULT 5
US-09-789-390-2
; Sequence 2, Application US/09789390
; GENERAL INFORMATION:
;  APPLICANT: Vernet, Corine
;  APPLICANT:  Fernandes, Elma
;  APPLICANT:  MacDougall, John
;  APPLICANT:  Shimkets, Richard A
;  APPLICANT:  Spaderna, Steven K
;  TITLE OF INVENTION: NOVEL POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
;  FILE REFERENCE: 15966-692
;  CURRENT APPLICATION NUMBER: US/09/789,390
;  CURRENT FILING DATE:  2001-02-23
;  PRIOR APPLICATION NUMBER: 60/185,548
;  PRIOR FILING DATE: 2000-02-28
;  PRIOR APPLICATION NUMBER: 60/199,957
;  PRIOR FILING DATE: 2000-04-27
;  PRIOR APPLICATION NUMBER: 60/184,951
;  PRIOR FILING DATE: 2000-02-25
;  PRIOR APPLICATION NUMBER: 60/185,967
;  PRIOR FILING DATE: 2000-[illegible]
;  PRIOR APPLICATION NUMBER: 60/197,723
;  PRIOR FILING DATE: 2000-04-18
;  NUMBER OF SEQ ID NOS: 77
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 1281
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-789-390-2

  Query Match             75.5%;  Score 7060;  DB 21;  Length 1281;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 1279;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        411 MPYDHFQPLPRWEHNPWTACSVSCGGGIQRRSFVCVEESMHGEILQVEEWKCMYAPKPKV 470
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MPYDHFQPLPRWEHNPWTACSVSCGGGIQRRSFVCVEESMHGEILQVEEWKCMYAPKPKV 60

Qy        471 MQTCNLFDCPKWIAMEWSQCTVTCGRGLRYRVVLCINHRGEHVGGCNPQLKLHIKEECVI 530
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 MQTCNLFDCPKWIAMEWSQCTVTCGRGLRYRVVLCINHRGEHVGGCNPQLKLHIKEECVI 120

Qy        531 PIPCYKPKEKSPVEAKLPWLKQAQELEETRIATEEPTFIPEPWSACSTTCGPGVQVREVK 590
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 PIPCYKPKEKSPVEAKLPWLKQAQELEETRIATEEPTFIPEPWSACSTTCGPGVQVREVK 180

Qy        591 CRVLLTFTQTETELPEEECEGPKLPTERPCLLEACDESPASPELDIPLPEDSETTYDWEY 650
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CRVLLTFTQTETELPEEECEGPKLPTERPCLLEACDESPASPELDIPLPEDSETTYDWEY 240

Qy        651 AGFTPCTATCLGGHQEAIAVCLHIQTQQTVNDSLCDMVHRPPAMSQACNTEPCPPRWHVG 710
              ||||||||||:|||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGFTPCTATCVGGHQEAIAVCLHIQTQQTVNDSLCDMVHRPPAMSQACNTEPCPPRWHVG 300

Qy        711 SWGPCSATCGVGIQTRDVYCLHPGETPAPPEECRDEKPHALQACNQFDCPPGWHIEEWQQ 770
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 SWGPCSATCGVGIQTRDVYCLHPGETPAPPEECRDEKPHALQACNQFDCPPGWHIEEWQQ 360

Qy        771 CSRTCGGGTQNRRVTCRQLLTDGSFLNLSDELCQGPKASSHKSCARTDCPPHLAVGDWSK 830
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CSRTCGGGTQNRRVTCRQLLTDGSFLNLSDELCQGPKASSHKSCARTDCPPHLAVGDWSK 420

Qy        831 CSVSCGVGIQRRKQVCQRLAAKGRRIPLSEMMCRDLPGEPLVRSCQMPECSKIKSEMKTK 890
              |||||||||||||||||||||||||||||||||| ||| |||||||||||||||||||||
Db        421 CSVSCGVGIQRRKQVCQRLAAKGRRIPLSEMMCRDLPGLPLVRSCQMPECSKIKSEMKTK 480

Qy        891 LGEQGPQILSVQRVYIQTREEKRINLTIGSPAYLLPNTSVIIKCPVRRFQKSLIQWEKDG 950
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 LGEQGPQILSVQRVYIQTREEKRINLTIGSPAYLLPNTSVIIKCPVRRFQKSLIQWEKDG 540

Qy        951 RCLQNSKRLGITKSGSLKIHGLAAPDIGVYRCIAGSAQETVVLKLIGTDNRLIARPALRE 1010
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 RCLQNSKRLGITKSGSLKIHGLAAPDIGVYRCIAGSAQETVVLKLIGTDNRLIARPALRE 600

Qy       1011 PMREYPGMDHSEANSLGVTWHKMRQMWNNKNDLYLDDDHISNQPFLRALLGHCSNSAGST 1070
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 PMREYPGMDHSEANSLGVTWHKMRQMWNNKNDLYLDDDHISNQPFLRALLGHCSNSAGST 660

Qy       1071 NSWELKNKQFEAAVKQGAYSMDTAQFDELIRNMSQLMETGEVSDDLASQLIYQLVAELAK 1130
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 NSWELKNKQFEAAVKQGAYSMDTAQFDELIRNMSQLMETGEVSDDLASQLIYQLVAELAK 720

Qy       1131 AQPTHMQWRGIQEETPPAAQLRGETGSVSQSSHAKNSGKLTFKPKGPVLMRQSQPPSISF 1190
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AQPTHMQWRGIQEETPPAAQLRGETGSVSQSSHAKNSGKLTFKPKGPVLMRQSQPPSISF 780

Qy       1191 NKTINSRIGNTVYITKRTFVINILCDLITPSEATYTWTKDGTLLQPSVKIILDGTGKIQI 1250
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 NKTINSRIGNTVYITKRTFVINILCDLITPSEATYTWTKDGTLLQPSVKIILDGTGKIQI 840

Qy       1251 QNPTRKEQGIYECSVANHLGSDVESSSVLYAEAPVILSVERNITKPEHNHLSVVVGGIVE 1310
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 QNPTRKEQGIYECSVANHLGSDVESSSVLYAEAPVILSVERNITKPEHNHLSVVVGGIVE 900

Qy       1311 AALGANVIIRCPVKGVPQPNITWLKRGGSLSGNVSLLFNGSLLLQNVSLENEGTYVCIAT 1370
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AALGANVIIRCPVKGVPQPNITWLKRGGSLSGNVSLLFNGSLLLQNVSLENEGTYVCIAT 960

Qy       1371 NALGKAVATSVLHLLERRWPESRIVFLQGHKKYILQATNTRTNSNDPTGEPPPQEPFWEP 1430
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 NALGKAVATSVLHLLERRWPESRIVFLQGHKKYILQATNTRTNSNDPTGEPPPQEPFWEP 1020

Qy       1431 GNWSHCSATCGHLGARIQRPQCVMANGQEVSEALCDHLQKPLAGFEPCNIRDCPARWFTS 1490
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GNWSHCSATCGHLGARIQRPQCVMANGQEVSEALCDHLQKPLAGFEPCNIRDCPARWFTS 1080

Qy       1491 VWSQCSVSCGEGYHSRQVTCKRTKANGTVQVVSPRACAPKDRPLGRKPCFGHPCVQWEPG 1550
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 VWSQCSVSCGEGYHSRQVTCKRTKANGTVQVVSPRACAPKDRPLGRKPCFGHPCVQWEPG 1140

Qy       1551 NRCPGRCMGRAVRMQQRHTACQHNSSDSNCDDRKRPTLRRNCTSGACDVCWHTGPWKPCT 1610
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 NRCPGRCMGRAVRMQQRHTACQHNSSDSNCDDRKRPTLRRNCTSGACDVCWHTGPWKPCT 1200

Qy       1611 AACGRGFQSRKVDCIHTRSCKPVAKRHCVQKKKPISWRHCLGPSCDRDCIDTTHYCMFVK 1670
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 AACGRGFQSRKVDCIHTRSCKPVAKRHCVQKKKPISWRHCLGPSCDRDCIDTTHYCMFVK 1260

Qy       1671 HLNLCSLDRYKQRCCQSCQEG 1691
              |||||||||||||||||||||
Db       1261 HLNLCSLDRYKQRCCQSCQEG 1281


RESULT 6
US-09-789-390-61
; Sequence 61, Application US/09789390
; GENERAL INFORMATION:
;  APPLICANT: Vernet, Corine
;  APPLICANT:  Fernandes, Elma
;  APPLICANT:  MacDougall, John
;  APPLICANT:  Shimkets, Richard A
;  APPLICANT:  Spaderna, Steven K
;  TITLE OF INVENTION: NOVEL POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
;  FILE REFERENCE: 15966-692
;  CURRENT APPLICATION NUMBER: US/09/789,390
;  CURRENT FILING DATE:  2001-02-23
;  PRIOR APPLICATION NUMBER: 60/185,548
;  PRIOR FILING DATE: 2000-02-28
;  PRIOR APPLICATION NUMBER: 60/199,957
;  PRIOR FILING DATE: 2000-04-27
;  PRIOR APPLICATION NUMBER: 60/184,951
;  PRIOR FILING DATE: 2000-02-25
;  PRIOR APPLICATION NUMBER: 60/185,967
;  PRIOR FILING DATE: 2000-03-01
;  PRIOR APPLICATION NUMBER: 60/197,723
;  PRIOR FILING DATE: 2000-04-18
;  NUMBER OF SEQ ID NOS: 77
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 61
;   LENGTH: 1023
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-789-390-61

  Query Match             59.9%;  Score 5601;  DB 21;  Length 1023;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1022;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        669 AVCLHIQTQQTVNDSLCDMVHRPPAMSQACNTEPCPPRWHVGSWGPCSATCGVGIQTRDV 728
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AVCLHIQTQQTVNDSLCDMVHRPPAMSQACNTEPCPPRWHVGSWGPCSATCGVGIQTRDV 60

Qy        729 YCLHPGETPAPPEECRDEKPHALQACNQFDCPPGWHIEEWQQCSRTCGGGTQNRRVTCRQ 788
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 YCLHPGETPAPPEECRDEKPHALQACNQFDCPPGWHIEEWQQCSRTCGGGTQNRRVTCRQ 120

Qy        789 LLTDGSFLNLSDELCQGPKASSHKSCARTDCPPHLAVGDWSKCSVSCGVGIQRRKQVCQR 848
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 LLTDGSFLNLSDELCQGPKASSHKSCARTDCPPHLAVGDWSKCSVSCGVGIQRRKQVCQR 180

Qy        849 LAAKGRRIPLSEMMCRDLPGEPLVRSCQMPECSKIKSEMKTKLGEQGPQILSVQRVYIQT 908
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 LAAKGRRIPLSEMMCRDLPGEPLVRSCQMPECSKIKSEMKTKLGEQGPQILSVQRVYIQT 240

Qy        909 REEKRINLTIGSPAYLLPNTSVIIKCPVRRFQKSLIQWEKDGRCLQNSKRLGITKSGSLK 968
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 REEKRINLTIGSPAYLLPNTSVIIKCPVRRFQKSLIQWEKDGRCLQNSKRLGITKSGSLK 300

Qy        969 IHGLAAPDIGVYRCIAGSAQETVVLKLIGTDNRLIARPALREPMREYPGMDHSEANSLGV 1028
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 IHGLAAPDIGVYRCIAGSAQETVVLKLIGTDNRLIARPALREPMREYPGMDHSEANSLGV 360

Qy       1029 TWHKMRQMWNNKNDLYLDDDHISNQPFLRALLGHCSNSAGSTNSWELKNKQFEAAVKQGA 1088
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TWHKMRQMWNNKNDLYLDDDHISNQPFLRALLGHCSNSAGSTNSWELKNKQFEAAVKQGA 420

Qy       1089 YSMDTAQFDELIRNMSQLMETGEVSDDLASQLIYQLVAELAKAQPTHMQWRGIQEETPPA 1148
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 YSMDTAQFDELIRNMSQLMETGEVSDDLASQLIYQLVAELAKAQPTHMQWRGIQEETPPA 480
```

```
Qy       1149 AQLRGETGSVSQSSHAKNSGKLTFKPKGPVLMRQSQPPSISFNKTINSRIGNTVYITKRT 1208
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AQLRGETGSVSQSSHAKNSGKLTFKPKGPVLMRQSQPPSISFNKTINSRIGNTVYITKRT 540

Qy       1209 EVINILCDLITPSEATYTWTKDGTLLQPSVKIILDGTGKIQIQNPTRKEQGIYECSVANH 1268
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 EVINILCDLITPSEATYTWTKDGTLLQPSVKIILDGTGKIQIQNPTRKEQGIYECSVANH 600

Qy       1269 LGSDVESSSVLYAEAPVILSVERNITKPEHNHLSVVVGGIVEAALGANVTIRCPVKGVPQ 1328
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 LGSDVESSSVLYAEAPVILSVERNITKPEHNHLSVVVGGIVEAALGANVTIRCPVKGVPQ 660

Qy       1329 PNITWLKRGGSLSGNVSLLFNGSLLLQNVSLENEGTYVCIATNALGKAVATSVLHLLERR 1388
              ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db        661 PNITWLKRGGSLSGNVSLLFNGSLLLQNVSLENEGTYVCIATNALGKAVATSVFHLLERR 720

Qy       1389 WPESRTVPLQGHKKYILQATNTRTNSNDPTGEPPPQEPFWEPGNWSHCSATCGHLGARIQ 1448
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 WPESRTVPLQGHKKYILQATNTRTNSNDPTGEPPPQEPFWEPGNWSHCSATCGHLGARIQ 780

Qy       1449 RPQCVMANGQEVSEALCDHLQKPLAGFEPCNIRDCPARWFTSVWSQCSVSCGEGYHSRQV 1508
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 RPQCVMANGQEVSEALCDHLQKPLAGFEPCNIRDCPARWFTSVWSQCSVSCGEGYHSRQV 840

Qy       1509 TCKRTKANGTVQVVSPRACAPKDRPLGRKPCFGHPCVQWEPGNRCPGRCMGRAVRMQQRH 1568
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TCKRTKANGTVQVVSPRACAPKDRPLGRKPCFGHPCVQWEPGNRCPGRCMGRAVRMQQRH 900

Qy       1569 TACQHNSSDSNCDDRKRPTLRRNCTSGACDVCWHTGPWKPCTAACGRGFQSRKVDCIHTR 1628
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TACQHNSSDSNCDDRKRPTLRRNCTSGACDVCWHTGPWKPCTAACGRGFQSRKVDCIHTR 960

Qy       1629 SCKPVAKRHCVQKKKPISWRHCTQPSCDRDCTDTTHYCMFVKHLNLCSLDRYKQRCCQSC 1688
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 SCKPVAKRHCVQKKKPISWRHCTQPSCDRDCTDTTHYCMFVKHLNLCSLDRYKQRCCQSC 1020

Qy       1689 QEG 1691
              |||
Db       1021 QEG 1023


RESULT 7
US-09-789-390-25
; Sequence 25, Application US/09789390
; GENERAL INFORMATION:
;  APPLICANT: Vernet, Corine
;  APPLICANT:  Fernandes, Elma
;  APPLICANT:  MacDougall, John
;  APPLICANT:  Shimkets, Richard A
;  APPLICANT:  Spaderna, Steven K
;  TITLE OF INVENTION: NOVEL POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
;  FILE REFERENCE: 15966-652
;  CURRENT APPLICATION NUMBER: US/09/789,390
;  CURRENT FILING DATE:  2001-02-23
;  PRIOR APPLICATION NUMBER: 60/185,548
;  PRIOR FILING DATE: 2000-02-28
;  PRIOR APPLICATION NUMBER: 60/199,957
;  PRIOR FILING DATE: 2000-04-27
;  PRIOR APPLICATION NUMBER: 60/184,951
;  PRIOR FILING DATE: 2000-02-25
;  PRIOR APPLICATION NUMBER: 60/185,967
;  PRIOR FILING DATE: 2000-03-01
;  PRIOR APPLICATION NUMBER: 60/197,723
;  PRIOR FILING DATE: 2000-04-18
;  NUMBER OF SEQ ID NOS: 77
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 25
;    LENGTH: 1023
;    TYPE: PRT
;    ORGANISM: Homo sapiens
US-09-789-390-25

  Query Match             59.9%;  Score 5599;  DB 21;  Length 1023;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1022;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        669 AVCLHIQTQQIVNESGCDMVHRFPAMSQACNIEPCPPRWHVGSWGPCSATCGVGIQTRDV 728
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AVCLHIQTQQIVNESGCDMVHRFPAMSQACNIEPCPPRWHVGSWGPCSATCGVGIQTRDV 60

Qy        729 YCLHPGETPAPPEECRDEKPHALQACNQFDCPPGWHIEEWQQCSRTCGGGTQNRRVTCRQ 788
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 YCLHPGETPAPPEECRDEKPHALQACNQFDCPPGWHIEEWQQCSRTCGGGTQNRRVTCRQ 120

Qy        789 LLTDGSFLNLSDELCQGPKASSHKSCARTDCPPHLAVGDWSKCSVSCGVGLQRRKQVCQR 848
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 LLTDGSFLNLSDELCQGPKASSHKSCARTDCPPHLAVGDWSKCSVSCGVGLQRRKQVCQR 180

Qy        849 LAAKGRRIPLSEMMCRDLPGFPLVRSCQMPECSKIKSEMKTKLGEQGPQILSVQRVYIQT 908
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 LAAKGRRIPLSEMMCRDLPGFPLVRSCQMPECSKIKSEMKTKLGEQGPQILSVQRVYIQT 240

Qy        909 REEKRINLTIGSRAYLLPNTSVIIKCPVRRFQKSLIQWEKDGRCLQNSKRLGITKSGSLK 968
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 REEKRINLTIGSRAYLLPNTSVIIKCPVRRFQKSLIQWEKDGRCLQNSKRLGITKSGSLK 300

Qy        969 IHGLAAPDIGVYRCIAGSAQETVVLKLIGTDNRLIARPALREPMREYPGMDHSEANSLGV 1028
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 IHGLAAPDIGVYRCIAGSAQETVVLKLIGTDNRLIARPALREPMREYPGMDHSEANSLGV 360

Qy       1029 TWHKMRQMWNNKNDLYLDDDHISNQPFLRALLGHCSNSAGSTNSWELKNKQFEAAVKQGA 1088
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TWHKMRQMWNNKNDLYLDDDHISNQPFLRALLGHCSNSAGSTNSWELKNKQFEAAVKQGA 420

Qy       1089 YSMDTAQFDELIRNMSQLMETGEVSDDLASQLIYQLVAELAKAQPTHMQWRGIQEETPPA 1148
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 YSMDTAQFDELIRNMSQLMETGEVSDDLASQLIYQLVAELAKAQPTHMQWRGIQEETPPA 480

Qy       1149 AQLRGETGSVSQSSHAKNSGKLTFKPKGPVLMRQSQPPSISFNKTINSRIGNTVYITKRT 1208
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AQLRGETGSVSQSSHAKNSGKLTFKPKGPVLMRQSQPPSISFNKTINSRIGNTVYITKRT 540

Qy       1209 EVINILCDLITPSEATYTWTKDGTLLQPSVKIILDGTGKIQIQNPTRKEQGIYECSVANH 1268
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 EVINILCDLITPSEATYTWTKDGTLLQPSVKIILDGTGKIQIQNPTRKEQGIYECSVANH 600

Qy       1269 LGSDVESSSVLYAEAPVILSVERNITKPEHNHLSVVVGGIVEAALGANVTIRCPVKGVPQ 1328
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 LGSDVESSSVLYAEAPVILSVERNITKPEHNHLSVVVGGIVEAALGANVTIRCPVKGVPQ 660

Qy       1329 PNITWLKRGGSLSGNVSLLFNGSLLLQNVSLENEGTYVCIATNALGKAVATSVLHLLERR 1388
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 PNITWLKRGGSLSGNVSLLFNGSLLLQNVSLENEGTYVCIATNALGKAVATSVLHLLERR 720

Qy       1389 WPESRTVPLQGHKKYILQATNTRTNSNDPTGEPPPQEPFWEPGNWSHCSATCGHLGARIQ 1448
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 WPESRTVPLQGHKKYILQATNTRTNSNDPTGEPPPQEPFWEPGNWSHCSATCGHLGARIQ 780

Qy       1449 RPQCVMANGQEVSEALCDHLQKPLAGFEPCNIRDCPARWFTSVWSQCSVSCGEGYHSRQV 1508
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 RPQCVMANGQEVSEALCDHLQKPLAGFEPCNIRDCPARWFTSVWSQCSVSCGEGYHSRQV 840

Qy       1509 TCKRTKANGTVQVVSPRACAPKDRPLGRKPCFGHPCVQWEPGNRCPGRCMGRAVRMQQRH 1568
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TCKRTKANGTVQVVSPRACAPKDRPLGRKPCFGHPCVQWEPGNRCPGRCMGRAVRMQQRH 900

Qy       1569 TACQHNSSDSNCDDRKRPTLRRNCTSGACDVCWHTGPWKPCTAACGRGFQSRKVDCIHTR 1628
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TACQHNSSDSNCDDRKRPTLRRNCTSGACDVCWHTGPWKPCTAACGRGFQSRKVDCIHTR 960

Qy       1629 SCKPVAKRHCVQKKKPISWRHCTQPSCDRDCTDTTHYCMFVKHLNLCSLDRYKQRCCQSC 1688
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 SCKPVAKRHCVQKKKPISWRHCTQPSCDRDCTDTTHYCMFVKHLNLCSLDRYKQRCCQSC 1020

Qy       1689 QEG 1691
              |||
Db   1021 QEG 1023


RESULT 8
US-09-789-390-27
; Sequence 27, Application US/09789390
; GENERAL INFORMATION:
;  APPLICANT: Vernet, Corine
;  APPLICANT:  Fernandes, Elma
;  APPLICANT:  MacDougall, John
;  APPLICANT:  Shimkets, Richard A
;  APPLICANT:  Spaderna, Steven K
;  TITLE OF INVENTION. NOVEL POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
;  FILE REFERENCE: 15966-692
;  CURRENT APPLICATION NUMBER. US/09/789,390
;  CURRENT FILING DATE:  2001-02-23
;  PRIOR APPLICATION NUMBER: 60/185,548
;  PRIOR FILING DATE: 2000-02-28
;  PRIOR APPLICATION NUMBER: 60/199,957
;  PRIOR FILING DATE: 2000-04-27
;  PRIOR APPLICATION NUMBER: 60/184,951
;  PRIOR FILING DATE: 2000-02-25
;  PRIOR APPLICATION NUMBER: 60/185,967
;  PRIOR FILING DATE: 2000-03-01
;  PRIOR APPLICATION NUMBER: 60/197,723
;  PRIOR FILING DATE: 2000-04-18
;  NUMBER OF SEQ ID NOS. 77
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 27
;   LENGTH: 1023
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-789-390-27

  Query Match             59.9%;  Score 5599;  DB 21;  Length 1023;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1022,  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        669 AVCLHIQTQQTVNDSLCDMVHRPPAMSQACNTEPCPPRWHVGSWGPCSATCGVGIQTRDV 728
              ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db          1 AVCLHIQTQQTVNDSLCDMVHRPPAMSQACNTEPCPPRWHVGSWGPCSATCGVGIQTRDV 60

Qy        729 YCLHPGETPAPPEECRDEKPHALQACNQFDCPPGWHIEEWQQCSRTCGGGTQNRRVTCRQ 788
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 YCLHPGETPAPPEECRDEKPHALQACNQFDCPPGWHIEEWQQCSRTCGGGTQNRRVTCRQ 120

Qy        789 LLTDGSFLNLSDELCQGPKASSHKSCARTDCPPHLAVGDWSKCSVSCGVGIQRRKQVCQR 848
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 LLTDGSFLNLSDELCQGPKASSHKSCARTDCPPHLAVGDWSKCSVSCGVGIQRRKQVCQR 180

Qy        849 LAAKCRRIPLSEMMCRDLPGLPLVRSCQMPECSKIKSEMKIKLGEQGPQILSVQRVYIQT 908
              |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db        181 LAAKCRRIPLSEMMCRDLPGLPLVRSCQMPECSKIKSEMKIKLGEQGPQILSVQRVYIQT 240

Qy        909 REEKRINLTIGSRAYLLPNTSVIIKCPVRRFQKSLIQWEKDGRCLQNSKRLGITKSGSLK 968
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 REEKRINLTIGSRAYLLPNTSVIIKCPVRRFQKSLIQWEKDGRCLQNSKRLGITKSGSLK 300

Qy        969 IHGLAAPDIGVYRCIAGSAQETVVLKLIGTDNRLIARPALREPMREYPGMDHSEANSLGV 1028
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 IHGLAAPDIGVYRCIAGSAQETVVLKLIGTDNRLIARPALREPMREYPGMDHSEANSLGV 360

Qy       1029 TWHKMRQMWNNKNDLYLDDDHISNQPFLRALLGHCSNSAGSTNSWELKNKQFEAAVKQGA 1088
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TWHKMRQMWNNKNDLYLDDDHISNQPFLRALLGHCSNSAGSTNSWELKNKQFEAAVKQGA 420

Qy       1089 YSMDTAQFDELIRNMSQLMETGEVSDDLASQLIYQLVAELAKAQPTHMQWRGIQEETPPA 1148
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 YSMDTAQFDELIRNMSQLMETGEVSDDLASQLIYQLVAELAKAQPTHMQWRGIQEETPPA 480

Qy       1149 AQLRGETGSVSQSSHAKNSGKLTFKPKGPVLMRQSQPPSISFNKTINSRIGNIVYIIKRT 1208
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AQLRGETGSVSQSSHAKNSGKLTFKPKGPVLMRQSQPPSISFNKTINSRIGNIVYIIKRT 540
```

```
Qy       1209 EVINILCDLITPSEATYTWTKDGTLLQPSVKIILDGTGKIQIQNPTRKEQGIYECSVANH 1268
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 EVINILCDLITPSEATYTWTKDGTLLQPSVKIILDGTGKIQIQNPTRKEQGIYECSVANH 600

Qy       1269 LGSDVESSSVLYAEAPVILSVERNITKPEHNHLSVVVGGIVEAALGANVTIRCPVKGVPQ 1328
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 LGSDVESSSVLYAEAPVILSVERNITKPEHNHLSVVVGGIVEAALGANVTIRCPVKGVPQ 660

Qy       1329 PNITWLKRGGSLSGNVSLLFNGSLLLQNVSLENEGTYVCIATNALGKAVATSVLHLLERR 1388
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 PNITWLKRGGSLSGNVSLLFNGSLLLQNVSLENEGTYVCIATNALGKAVATSVLHLLERR 720

Qy       1389 WPESRIVFLQGHKKYILQATNTRTNSNDPTGEPPPQEPFWEPGNWSHCSATCGHLGARIQ 1448
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 WPESRIVFLQGHKKYILQATNTRTNSNDPTGEPPPQEPFWEPGNWSHCSATCGHLGARIQ 780

Qy       1449 RPQCVMANGQEVSEALCDHLQKPLAGFEPCNIRDCPARWFTSVWSQCSVSCGEGYHSRQV 1508
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 RPQCVMANGQEVSEALCDHLQKPLAGFEPCNIRDCPARWFTSVWSQCSVSCGEGYHSRQV 840

Qy       1509 TCKRTKANGTVQVVSPRACAPKDRPLGRKPCFGHPCVQWEPGNRCPGRCMGRAVRMQQRH 1568
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TCKRTKANGTVQVVSPRACAPKDRPLGRKPCFGHPCVQWEPGNRCPGRCMGRAVRMQQRH 900

Qy       1569 TACQHNSSDSNCDDRKRPTLRRNCTSGACDVCWHTGPWKPCTAACGRGFQSRKVDCIHTR 1628
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TACQHNSSDSNCDDRKRPTLRRNCTSGACDVCWHTGPWKPCTAACGRGFQSRKVDCIHTR 960

Qy       1629 SCKPVAKRHCVQKKKPISWRHCLGPSCDRDCTDTTHYCMFVKHLNLCSLDRYKQRCCQSC 1688
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 SCKPVAKRHCVQKKKPISWRHCLGPSCDRDCTDTTHYCMFVKHLNLCSLDRYKQRCCQSC 1020

Qy       1689 QEG 1691
              |||
Db       1021 QEG 1023


RESULT 9
US-09-789-390-28
; Sequence 28, Application US/09789390
; GENERAL INFORMATION:
;  APPLICANT: Vernet, Corine
;  APPLICANT:  Fernandes, Elma
;  APPLICANT:  MacDougall, John
;  APPLICANT:  Shimkets, Richard A
;  APPLICANT.  Spaderna, Steven K
;  TITLE OF INVENTION: NOVEL POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
;  FILE REFERENCE: 15966-692
;  CURRENT APPLICATION NUMBER: US/09/789,390
;  CURRENT FILING DATE:  2001-02-23
;  PRIOR APPLICATION NUMBER: 60/185,548
;  PRIOR FILING DATE. 2000-02-28
;  PRIOR APPLICATION NUMBER: 60/199,957
;  PRIOR FILING DATE: 2000-04-27
;  PRIOR APPLICATION NUMBER: 60/184,951
;  PRIOR FILING DATE: 2000-02-25
;  PRIOR APPLICATION NUMBER: 60/185,967
;  PRIOR FILING DATE: 2000-03-01
;  PRIOR APPLICATION NUMBER: 60/197,723
;  PRIOR FILING DATE: 2000-04-18
;  NUMBER OF SEQ ID NOS: 77
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 28
;   LENGTH: 1023
;   TYPE: PRT
;   ORGANISM: homo sapiens
US-09-789-390-28

  Query Match             59.9%;  Score 5599;  DB 21;  Length 1023;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1022;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;
```

```
Qy        669 AVCLHIQTQQTVNDSLCDMVHRPPAMSQACNTEPCPPRWHVGSWGPCSATCGVGIQTRDV 728
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AVCLHIQTQQTVNDSLCDMVHRPPAMSQACNTEPCPPRWHVGSWGPCSATCGVGIQTRDV 60

Qy        729 YCLHPGEIPAPPEECRDEKPHALQACNQFDCPPGWHIEEWQQCSRTCGGGIQNRRVTCRQ 788
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 YCLHPGEIPAPPEECRDEKPHALQACNQFDCPPGWHIEEWQQCSRTCGGGIQNRRVTCRQ 120

Qy        789 LLTDGSFLNLSDELCQGPKASSHKSCARTDCPPHLAVGDWSKCSVSCGVGIQRRKQVCQR 848
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 LLTDGSFLNLSDELCQGPKASSHKSCARTDCPPHLAVGDWSKCSVSCGVGIQRRKQVCQR 180

Qy        849 LAAKGRRIPLSEMMCRDLPGFPLVRSCQMPECSKIKSEMKTKLGEQGPQILSVQRVYIQT 908
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 LAAKGRRIPLSEMMCRDLPGFPLVRSCQMPECSKIKSEMKTKLGEQGPQILSVQRVYIQT 240

Qy        909 REEKRINLTIGSRAYLLPNTSVIIKCPVRRFQKSLIQWEKDGRCLQNSKRLGITKSGSLK 968
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 REEKRINLTIGSRAYLLPNTSVIIKCPVRRFQKSLIQWEKDGRCLQNSKRLGITKSGSLK 300

Qy        969 IHGLAAPDIGVYRCIAGSAQEIVVLKLIGTDNRLIARPALREPMREYPGMDHSEANSLGV 1028
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 IHGLAAPDIGVYRCIAGSAQEIVVLKLIGTDNRLIARPALREPMREYPGMDHSEANSLGV 360

Qy       1029 TWHKMRQMWNNKNDLYLDDDHISNQPFLRALLGHCSNSAGSTNSWELKNKQFEAAVKQGA 1088
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TWHKMRQMWNNKNDLYLDDDHISNQPFLRALLGHCSNSAGSTNSWELKNKQFEAAVKQGA 420

Qy       1089 YSMDTAQFDELIRNMSQLMETGEVSDDLASQLIYQLVAELAKAQPTHMQWRGIQEEIPPA 1148
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 YSMDTAQFDELIRNMSQLMETGEVSDDLASQLIYQLVAELAKAQPTHMQWRGIQEEIPPA 480

Qy       1149 AQLRGETGSVSQSSHAKNSGKLTFKPKGPVLMRQSQPPSISFNKTINSRIGNTVYITKRT 1208
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AQLRGETGSVSQSSHAKNSGKLTFKPKGPVLMRQSQPPSISFNKTINSRIGNTVYITKRT 540

Qy       1209 EVINILCDLITPSEATYTWTKDGTLLQPSVKIILDGTGKIQIQNPTRKEQGIYECSVANH 1268
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 EVINILCDLITPSEATYTWIKDGTLLQPSVKIILDGTGKIQIQNPTRKEQGIYECSVANH 600

Qy       1269 LGSDVESSSVLYAEAPVILSVERNITKPEHNHLSVVVGGIVEAALGANVTIRCPVKGVPQ 1328
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 LGSDVESSSVLYAEAPVILSVERNITKPEHNHLSVVVGGIVEAALGANVTIRCPVKGVPQ 660

Qy       1329 PNITWLKRGGSLSGNVSLLFNGSLLLQNVSLENEGTYVCIATNALGKAVATSVLHLLERR 1388
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 PNITWLKRGGSLSGNVSLLFNGSLLLQNVSLENEGTYVCIATNALGKAVATSVLHLLERR 720

Qy       1389 WPESRIVFLQGHKKYILQATNTRTNSNDPTGEPPPQEPFWEPGNWSHCSATCGHLGARIQ 1448
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 WPESRIVFLQGHKKYILQATNTRTNSNDPTGEPPPQEPFWEPGNWSHCSATCGHLGARIQ 780

Qy       1449 RPQCVMANGQEVSEALCDHLQKPLAGFEPCNIRDCPARWFTSVWSQCSVSCGEGYHSRQV 1508
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 RPQCVMANGQEVSEALCDHLQKPLAGFEPCNIRDCPARWFTSVWSQCSVSCGEGYHSRQV 840

Qy       1509 TCKRTKANGTVQVVSPRACAPKDRPLGRKPCFGHPCVQWEPGNRCPGRCMGRAVRMQQRH 1568
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TCKRTKANGTVQVVSPRACAPKDRPLGRKPCFGHPCVQWEPGNRCPGRCMGRAVRMQQRH 900

Qy       1569 TACQHNSGDSNCDDRKRPTLRRNCTSGACDVCWHTGPWKPCTAACGRGFQSRKVDCIHTR 1628
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TACQHNSGDSNCDDRKRPTLRRNCTSGACDVCWHTGPWKPCTAACGRGFQSRKVDCIHTR 960

Qy       1629 SCKPVAKRHCVQKKKPISWRHCLGPSCDRDCTDTTHYCMFVKHLNLCSLDRYKQRCCQSC 1688
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 SCKPVAKRHCVQKKKPISWRHCLGPSCDRDCTDTTHYCMFVKHLNLCSLDRYKQRCCQSC 1020

Qy       1689 QEG 1691
              |||
Db       1021 QEG 1023
```

```
RESULT 16
US-09-789-390-59
; Sequence 59, Application US/09789390
; GENERAL INFORMATION:
;  APPLICANT: Vernet, Corine
;  APPLICANT:  Fernandes, Elma
;  APPLICANT:  MacDougall, John
;  APPLICANT:  Shimkets, Richard A
;  APPLICANT:  Spaderna, Steven K
;  TITLE OF INVENTION: NOVEL POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
;  FILE REFERENCE: 15966-692
;  CURRENT APPLICATION NUMBER: US/09/789,390
;  CURRENT FILING DATE:  2001-02-23
;  PRIOR APPLICATION NUMBER: 60/185,548
;  PRIOR FILING DATE: 2000-02-28
;  PRIOR APPLICATION NUMBER: 60/199,957
;  PRIOR FILING DATE: 2000-04-27
;  PRIOR APPLICATION NUMBER: 60/184,951
;  PRIOR FILING DATE: 2000-02-25
;  PRIOR APPLICATION NUMBER: 60/185,967
;  PRIOR FILING DATE: 2000-03-01
;  PRIOR APPLICATION NUMBER: 60/197,723
;  PRIOR FILING DATE: 2000-04-18
;  NUMBER OF SEQ ID NOS: 77
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 59
;   LENGTH: 1023
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-789-390-59

  Query Match             59.9%;  Score 5599;  DB 21;  Length 1023;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1022;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        669 AVCLHIQTQQTVNDSLCDMVHRPPAMSQACNTEPCPPRWHVGSWGPCSATCGVGIQTRDV 728
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AVCLHIQTQQTVNDSLCDMVHRPPAMSQACNTEPCPPRWHVGSWGPCSATCGVGIQTRDV 60

Qy        729 YCLHPGEIPAPPEECRDEKPHALQACNQFDCPPGWHIEEWQQCSRTCGGGIQNRRVTCRQ 788
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 YCLHPGEIPAPPEECRDEKPHALQACNQFDCPPGWHIEEWQQCSRTCGGGIQNRRVTCRQ 120

Qy        789 LLTDGSFLNLSDELCQGPKASSHKSCARTDCPPHLAVGDWSKCSVSCGVGIQRRKQVCQR 848
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 LLTDGSFLNLSDELCQGPKASSHKSCARTDCPPHLAVGDWSKCSVSCGVGIQRRKQVCQR 180

Qy        849 LAAKGRRIPLSEMMCRDLPGFPLVRSCQMPECSKIKSEMKTKLGEQGPQILSVQRVYIQT 908
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 LAAKGRRIPLSEMMCRDLPGFPLVRSCQMPECSKIKSEMKTKLGEQGPQILSVQRVYIQT 240

Qy        909 REEKRINLTIGSRAYLLPNTSVIIKCPVRRFQKSLIQWEKDGRCLQNSKRLGITKSGSLK 968
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 REEKRINLTIGSRAYLLPNTSVIIKCPVRRFQKSLIQWEKDGRCLQNSKRLGITKSGSLK 300

Qy        969 IHGLAAPDIGVYRCIAGSAQEIVVLKLIGTDNRLIARPALREPMREYPGMDHSEANSLGV 1028
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 IHGLAAPDIGVYRCIAGSAQEIVVLKLIGTDNRLIARPALREPMREYPGMDHSEANSLGV 360

Qy       1029 TWHKMRQMWNNKNDLYLDDDHISNQPFLRALLGHCSNSAGSTNSWELKNKQFEAAVKQGA 1088
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TWHKMRQMWNNKNDLYLDDDHISNQPFLRALLGHCSNSAGSTNSWELKNKQFEAAVKQGA 420

Qy       1089 YSMDTAQFDELIRNMSQLMETGEVSDDLASQLIYQLVAELAKAQPTHMQWRGIQEEIPPA 1148
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 YSMDTAQFDELIRNMSQLMETGEVSDDLASQLIYQLVAELAKAQPTHMQWRGIQEEIPPA 480

Qy       1149 AQLRGETGSVSQSSHAKNSGKLTFKPKGPVLMRQSQPPSISFNKTINSRIGNTVYITKRT 1208
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AQLRGETGSVSQSSHAKNSGKLTFKPKGPVLMRQSQPPSISFNKTINSRIGNTVYITKRT 540

Qy       1209 EVINILCDLITPSEATYTWTKDGTLLQPSVKIILDGTGKIQIQNPTRKEQGIYECSVANH 1268
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      541 EVINILCDLITPSEATYIWTKDGTLLQPSVKIILDGTGKIQIQNPTRKEQGIYECSVANH 600

Qy     1269 LGSDVESSSVLYAEAPVILSVERNITKPEHNHLSVVVGGIVEAALGANVTIRCPVKGVPQ 1328
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      601 LGSDVESSSVLYAEAPVILSVERNITKPEHNHLSVVVGGIVEAALGANVTIRCPVKGVPQ 660

Qy     1329 PNITWLKRGGSLSGNVSLLFNGSLLLQNVSLENEGTYVCIATNALGKAVATSVLHLLERR 1388
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      661 PNITWLKRGGSLSGNVSLLFNGSLLLQNVSLENEGTYVCIATNALGKAVATSVLHLLERR 720

Qy     1389 WPESRIVPLQGHKKYILQATNTRTNSNDPTGEPPPQEPFWEPGNWSHCSATCGHLGARIQ 1448
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      721 WPESRIVPLQGHKKYILQATNTRTNSNDPTGEPPPQEPFWEPGNWSHCSATCGHLGARIQ 780

Qy     1449 RPQCVMANGQEVSEALCDHLQKPLAGFEPCNIRDCPARWFTSVWSQCSVSCGEGYHSRQV 1508
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      781 RPQCVMANGQEVSEALCDHLQKPLAGFEPCNIRDCPARWFTSVWSQCSVSCGEGYHSRQV 840

Qy     1509 TCKRTKANCIVQVVSPRACAPKDRPLGRKPCPGHPCVQWEPCNRCPCRCMCRAVRMQQRH 1568
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      841 TCKRTKANCIVQVVSPRACAPKDRPLGRKPCPGHPCVQWEPCNRCPCRCMCRAVRMQQRH 900

Qy     1569 TACQHNSSDSNCDDRKRPTLRRNCTSGACDVCWHTGPWKPCTAACGRGFQSRKVDCIHTR 1628
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      901 TACQHNSSDSNCDDRKRPTLRRNCTSGACDVCWHTGPWKPCTAACGRGFQSRKVDCIHTR 960

Qy     1629 SCKRVAKRHCVLKKKRISWRHCLGPSCDRDCTDTIHYCMFVKHGNLCSLPRYKQRCCQSC 1688
            |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db      961 SCKRVAKRHCVLKKKRISWRHCLGPSCDRDCIDIIHYCMFVKHGNLCSLPRYKQRCCQSC 1020

Qy     1689 QEG 1691
            |||
Db     1021 QEG 1023


RESULT 11
US-60-245-228-264
; Sequence 264, Application US/60245228
; GENERAL INFORMATION:
;  APPLICANT:  Beasley, Ellen
;  TITLE OF INVENTION:  ISOLATED HUMAN CYCLASE PROTEINS, NUCLEIC
;  TITLE OF INVENTION.  ACID MOLECULES ENCODING HUMAN CYCLASE PROTEINS, AND USES
;  TITLE OF INVENTION:  THEREOF
;  FILE REFERENCE:  CL000878
;  CURRENT APPLICATION NUMBER: US/60/245,228
;  CURRENT FILING DATE:  2000-11-03
;  NUMBER OF SEQ ID NOS:  630
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 264
;   LENGTH: 1036
;   TYPE: PRT
;   ORGANISM: HUMAN
US-60-245-228-264

  Query Match             57.5%;  Score 5372;  DB 27;  Length 1036;
  Best Local Similarity   70.6%;  Pred. No. 0;
  Matches 1034;  Conservative    0;  Mismatches    2;  Indels  428;  Gaps    4;

Qy      122 DCPPDAEDFRAQQCSAYNDVQYQGHYYEWLPRYNDPAAPCALKCHAQGQNLVVELAPKVL 181
            |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db        1 DCPPDAEDFRAQQCSAYNDVQYQGRYYEWLPRYNDPAAPCALKCHAQGQNLVVELAPKVL 60

Qy      182 DGTRCNIDSLDMCISGICQAVGCDRQLGSNAKEDNCGVCAGDGSTCRLVRGQSKSHVSPE 241
            ||||||||||||||||||
Db       61 DGTRCNIDSLDMCISGIC------------------------------------------ 78

Qy      242 KREENVIAVPLGSRSVRITVKGPAHLFIESKTLQGSKGEHSFNSPGVFVVENTTVEFQRG 301

Db       79 ------------------------------------------------------------ 78

Qy      302 SERQTFKIPGPLMADFIFKTRYTAAKDSVVQFFFYQPISHQWRQTDFFPCTVTCGGGYQL 361

Db       79 ------------------------------------------------------------ 78
```

```
Qy      362 NFAETVDIKLEEVVPCHTCHYYIENVEFPKLKQCSMLEQCRSLCFRKLMGYLHQRPK 421

Db       79 ------------------------------------------------------------ 78

Qy      422 WPHNPWTACSVSCGGGIQRRSFVCVEESMHGEILQVEEWKCMYAPKPKVMQTCNLFDCPK 481

Db       79 ------------------------------------------------------------ 78

Qy      482 WIAMEWSQCTVTCGRGLRYRVVLCINHRGEHVGGCNPQLKLHIKEECVIPIPCYKPKEKS 541

Db       79 ------------------------------------------------------------ 78

Qy      542 PVEAKLPWLKQAQELEETRIATEEPTFIPEPWSACSTTCGPGVQVREVKCRVLLTFTQTE 601
                                                       |||||||||||||||||
Db       79 -------------------------------------------QVREVKCRVLLTFTQTE 95

Qy      602 TELPEEECEGPKLPTERPCLLEACDESPASRELDIPLPEDSETTYDWEYAGFTPCTATCL 661
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       96 TELPEEECEGPKLPTERPCLLEACDESPASRELDIPLPEDSETTYDWEYAGFTPCTATCL 155

Qy      662 GGHQEAIAVCLHIQTQQTVNDSLCDMVHRPPAMSQACNTEPCPPRWHVGSWGPCSATCGV 721
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      156 GGHQEAIAVCLHIQTQQTVNDSLCDMVHRPPAMSQACNTEPCPPRWHVGSWGPCSATCGV 215

Qy      722 GIQTRDVYCLHPGETPAPPEECRDEKPHALQACNQFDCPPGWHIEEWQQCSRTCGGGIQN 781
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      216 GIQTRDVYCLHPGETPAPPEECRDEKPHALQACNQFDCPPGWHIEEWQQCSRTCGGGIQN 275

Qy      782 RRVTCRQLLTDGSFLNLSDELCQGPKASSHKSCARTDCPPHLAVGDWSKCSVSCGVGIQR 841
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      276 RRVTCRQLLTDGSFLNLSDELCQGPKASSHKSCARTDCPPHLAVGDWSKCSVSCGVGIQR 335

Qy      842 RKQVCQRLAAKGRRIPLSEMMCRDLPGFPLVRSCQMPECSKIKSEMKTKLGEQGPQILSV 901
            ||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db      336 RKQVCQRLAAKGRRIPLSEMMCRDLPGLPLVRSCQMPECSKIKSEMKTKLGEQGPQILSV 395

Qy      902 QRVYIQTREEKRINLTIGSRAYLLPNTSVIIKCPVRRFQKSLIQWEKDGRCLQNSKRLGI 961
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      396 QRVYIQTREEKRINLTIGSRAYLLPNTSVIIKCPVRRFQKSLIQWEKDGRCLQNSKRLGI 455

Qy      962 TKSGSLKIHGLAAPDIGVYRCIAGSAQETVVLKLIGTDNRLIARPALREPMREYPGMDHS 1021
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      456 TKSGSLKIHGLAAPDIGVYRCIAGSAQETVVLKLIGTDNRLIARPALREPMREYPGMDHS 515

Qy     1022 EANSLGVTWHKMRQMWNNKNDLYLDDDHISNQPELRALLGHTGNSAAYINSWELENKQFE 1081
            ||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||
Db      516 EANSLGVTWHKMRQMWNNKNDLYLDDDH--NQPELRALLGHTGNSAAYINSWELENKQFE 573

Qy     1082 AAVKQGAYSMDTAQFDELIRNMSQLMETGEVSDDLASQLIYQLVAELAKAQPTHMQWRGI 1141
            ||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||||
Db      574 AAVKQGAYSMDTAQFDELIRNMSQLM--GEVSDDLASQLIYQLVAELAKAQPTHMQWRGI 631

Qy     1142 QEETPPAAQLRGETGSVSQSSHAKNSGKLTFKPKGPVLMRQSQPPSISFNKTINSRIGNT 1201
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      632 QEETPPAAQLRGETGSVSQSSHAKNSGKLTFKPKGPVLMRQSQPPSISFNKTINSRIGNT 691

Qy     1202 VYITKRTEVINILCDLITPSEATYTWTKDGTLLQPSVKIILDGTGKIQIQNPTRKEQGIY 1261
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      692 VYITKRTEVINILCDLITPSEATYTWTKDGTLLQPSVKIILDGTGKIQIQNPTRKEQGIY 751

Qy     1262 ECSVANHLGSDVESSSVLYAEAPVILSVERNITKPEHNHLSVVVGGIVEAALGANVTIRC 1321
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      752 ECSVANHLGSDVESSSVLYAEAPVILSVERNITKPEHNHLSVVVGGIVEAALGANVTIRC 811

Qy     1322 PVKGVPQPNITWLKRGGSLSGNVSLLFNGSLLLQNVSLENEGTYVCIATNALGKAVATSV 1381
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      812 PVKGVPQPNITWLKRGGSLSGNVSLLFNGSLLLQNVSLENEGTYVCIATNALGKAVATSV 871

Qy     1382 LHLLERRWPESRIVPLQGHKKYILQATNTRTNSNDPTGEPPPQEPFWEPGNWSHCSATCG 1441
            ||||                                       |||||||||||||||||
Db      872 LHLL---------------------------------------EPFWEPGNWSHCSATCG 892
```

```
Qy       1442 HLGARIQRPQCVMANGQEVSEALCDHLQKPLAGFEPCNIRDCPARWFTSVWSQCSVSCGE 1501
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        893 HLGARIQRPQCVMANGQEVSEALCDHLQKPLAGFEPCNIRDCPARWFTSVWSQCSVSCGE 952

Qy       1502 GYHSRQVTCKRTKANGTVQVVSPRACAPKDRPLGRKPCFGHPCVQWEPGNRCPGRCMGRA 1561
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        953 GYHSRQVTCKRTKANGTVQVVSPRACAPKDRPLGRKPCFGHPCVQWEPGNRCPGRCMGRA 1012

Qy       1562 VRMQQRHTACQHNSSDSNCDDRKR 1585
              ||||||||||||||||||||||||
Db       1013 VRMQQRHTACQHNSSDSNCDDRKR 1036


RESULT 12
US-60-258-272-108
; Sequence 108, Application US/60258272
; GENERAL INFORMATION:
;  APPLICANT: Beasley, Ellen
;  TITLE OF INVENTION: ISOLATED HUMAN CYCLASE PROTEINS, NUCLEIC
;  TITLE OF INVENTION:  ACID MOLECULES ENCODING HUMAN CYCLASE PROTEINS, AND USES
;  TITLE OF INVENTION:  THEREOF
;  FILE REFERENCE: CL001041
;  CURRENT APPLICATION NUMBER: US/60/258,272
;  CURRENT FILING DATE:  2000-12-27
;  NUMBER OF SEQ ID NOS: 252
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 108
;   LENGTH: 1036
;   TYPE: PRT
;   ORGANISM: HUMAN
US-60-258-272-108

  Query Match             57.5%;  Score 5372;  DB 27;  Length 1036;
  Best Local Similarity   70.6%;  Pred. No. 0;
  Matches 1034;  Conservative    0;  Mismatches    2;  Indels  428;  Gaps    4;

Qy        122 DCPPDAEDFRAQQCSAYNDVQYQGRYYEWLPRYNDPAAPCALKCHAQGQNLVVELAPKVL 181
              ||||||||||||||||||||| || |||||||||||||||||||||||||||||||||||
Db          1 DCPPDAEDFRAQQCSAYNDVQYQGRYYEWLPRYNDPAAPCALKCHAQGQNLVVELAPKVL 60

Qy        182 DGIRCNIDSLDMCISGICQAVGCDRQLGSNAKEDNCGVCAGDGSTCRLVRGQSKSHVSPE 241
              ||||||||||||||||||
Db         61 DGIRCNIDSLDMCISGIC------------------------------------------ 78

Qy        242 KREENVIAVPLGSRSVRITVKGPAHLFIESKTLQGSKGEHSFNSPGVFVVENTTVEFQRG 301

Db         79 ------------------------------------------------------------ 78

Qy        302 SERQTFKIPGPLMADFIFKTRYTAAKDSVVQFFFYQPISHQWRQTDFFPCTVTCGGGYQL 361

Db         79 ------------------------------------------------------------ 78

Qy        362 NSAECVDIRLKRVVPDHYCHYYPENVKPKPKLKECSMDPCPSSDGFKEIMPYDHFQPLPR 421

Db         79 ------------------------------------------------------------ 78

Qy        422 WEHNPWTACSVSCGGGIQRRSFVCVEESMHGEILQVEEWKCMYAPKPKVMQTCNLFDCPK 481

Db         79 ------------------------------------------------------------ 78

Qy        482 WIAMEWSQCTVTCGRGLRYRVVLCINHRGEHVGGCNPQLKLHIKEECVIPIPCYKPKEKS 541

Db         79 ------------------------------------------------------------ 78

Qy        542 PVEAKLPWLKQAQELEETRIATEEPTFIPEPWSACSTTCGPGVQVREVKCRVLLTFTQTE 601
                                                         |||||||||||||||||
Db         79 -------------------------------------------QVREVKCRVLLTFTQTE 95

Qy        602 TELPEEECEGPKLPTERPCLLEACDESPASRELDIPLPEDSETTYDWEYAGFTPCTATCL 661
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         96 TELPEEECEGPKLPTERPCLLEACDESPASRELDIPLPEDSETTYDWEYAGFTPCTATCL 155

Qy        662 GGHQEATAVCLHIQTQQTVNDSLCDMVHRPPAMSQACNTEPCPPRWHVGSWGPCSATCGV 721
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        156 GGHQEATAVCLHIQTQQTVNDSLCDMVHRPPAMSQACNTEPCPPRWHVGSWGPCSATCGV 215

Qy        722 GIQTRDVYCLHPGETPAPPEECRDEKPHALQACNQFDCPPGWHIEEWQQCSRTCGGGTQN 781
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        216 GIQTRDVYCLHPGETPAPPEECRDEKPHALQACNQFDCPPGWHIEEWQQCSRTCGGGTQN 275

Qy        782 RRVTCRQLLTDGSFLNLSDELCQGPKASSHKSCARTDCPPHLAVGDWSKCSVSCGVGIQR 841
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        276 RRVTCRQLLTDGSFLNLSDELCQGPKASSHKSCARTDCPPHLAVGDWSKCSVSCGVGIQR 335

Qy        842 RKQVCQRLAAKGRRIPLSEMMCRDLPGFPLVRSCQMPECSKIKSEMKTKLGEQGPQILSV 901
              ||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db        336 RKQVCQRLAAKGRRIPLSEMMCRDLPGLPLVRSCQMPECSKIKSEMKTKLGEQGPQILSV 395

Qy        902 QRVYIQTREEKRINLTIGSRAYLLPNTSVIIKCPVRRFQKSLIQWEKDGRCLQNSKRLGI 961
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        396 QRVYIQTREEKRINLTIGSRAYLLPNTSVIIKCPVRRFQKSLIQWEKDGRCLQNSKRLGI 455

Qy        962 TKSGSLKIHGLAAPDIGVYRCIAGSAQETVVLKLIGTDNRLIARPALREPMREYPGMDHS 1021
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        456 TKSGSLKIHGLAAPDIGVYRCIAGSAQETVVLKLIGTDNRLIARPALREPMREYPGMDHS 515

Qy       1022 EANSLGVTWHKMRQMWNNKNDLYLDDDHISNQPFLRALLGHCSNSAGSTNSWELKNKQFE 1081
              ||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||
Db        516 EANSLGVTWHKMRQMWNNKNDLYLDDDH--NQPFLRALLGHCSNSAGSTNSWELKNKQFE 573

Qy       1082 AAVKQGAYSMDTAQFDELIRNMSQLMETGEVSDDLASQLIYQLVAELAKAQPTHMQWRGI 1141
              ||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||||
Db        574 AAVKQGAYSMDTAQFDELIRNMSQLM--GEVSGDLASQLIYQLVAELAKAQPTHMQWRGI 631

Qy       1142 QEETPPAAQLRGETGSVSQSSHAKNSGKLTFKPKGPVLMRQSQPPSISFNKTINSRIGNT 1201
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        632 QEETPPAAQLRGETGSVSQSSHAKNSGKLTFKPKGPVLMRQSQPPSISFNKTINSRIGNT 691

Qy       1202 VYIIKRTEVINILCDLITPSEATYTWTKDGTLLQPSVKIILDGTGKIQIQNPTRKEQGIY 1261
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        692 VYIIKRTEVINILCDLITPSEATYTWTKDGTLLQPSVKIILDGTGKIQIQNPTRKEQGIY 751

Qy       1262 ECSVANHLGSDVESSSVLYAEAPVILSVERNITKPEHNHLSVVVGGIVEAALGANVTIRC 1321
              |||||||||||||||||| | |||||||||||||||||||||||||||||||||||||||
Db        752 ECSVANHLGSDVESSSVLYAEAPVILSVERNITKPEHNHLSVVVGGIVEAALGANVTIRC 811

Qy       1322 PVKGVPQPNITWLKRGGSLSGNVSLLFNGSLLLQNVSLENEGTYVCIATNALGKAVATSV 1381
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        812 PVKGVPQPNITWLKRGGSLSGNVSLLFNGSLLLQNVSLENEGTYVCIATNALGKAVATSV 871

Qy       1382 LHLLERRWPESRIVPLQGHKKYILQATNTRTNSNDPTGEPPPQEPFWEPGNWSHCSATCG 1441
              ||||                                       |||||||||||||||||
Db        872 LHLL---------------------------------------EPFWEPGNWSHCSATCG 892

Qy       1442 HLGARIQRPQCVMANGQEVSEALCDHLQKPLAGFEPCNIRDCPARWFTSVWSQCSVSCGE 1501
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        893 HLGARIQRPQCVMANGQEVSEALCDHLQKPLAGFEPCNIRDCPARWFTSVWSQCSVSCGE 952

Qy       1502 GYHSRQVTCKRTKANGTVQVVSPRACAPKDRPLGRKPCFGHPCVQWEPGNRCPGRCMGRA 1561
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        953 GYHSRQVTCKRTKANGTVQVVSPRACAPKDRPLGRKPCFGHPCVQWEPGNRCPGRCMGRA 1012

Qy       1562 VRMQQRHTACQHNSSDSNCDDRKR 1585
              ||||||||||||||||||||||||
Db       1013 VRMQQRHTACQHNSSDSNCDDRKR 1036


RESULT 13
US-09-715-417A-28
; Sequence 28, Application US/09715417A
; GENERAL INFORMATION:
;  APPLICANT: Shimkets, Richard
;  APPLICANT:  Lichenstein, Henri
;  APPLICANT:  Vernet, Corine
;  APPLICANT:  Fernandes, Elma
;  TITLE OF INVENTION: NOVEL POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
;  FILE REFERENCE: 15966-606
;  CURRENT APPLICATION NUMBER: US/09/715,417A
;  CURRENT FILING DATE:  2001-11-16
;  PRIOR APPLICATION NUMBER: 60/166,336
;  PRIOR FILING DATE: 1999-11-19
;  PRIOR APPLICATION NUMBER: 60/167,785
;  PRIOR FILING DATE: 1999-11-29
;  PRIOR APPLICATION NUMBER: 60/187,844
;  PRIOR FILING DATE: 2000-03-08
;  NUMBER OF SEQ ID NOS: 38
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 28
;   LENGTH: 926
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-715-417A-28

  Query Match             54.4%;  Score 5086;  DB 21;  Length 926;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches  925;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        314 MADFIFKTRYTAAKDSVVQFFFYQPISHQWRQTDFFPCTVTCGGGYQLNSAECVDIRLKR 373
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MADFIFKTRYTAAKDSVVQFFFYQPISHQWRQTDFFPCTVTCGGGYQLNSAECVDIRLKR 60

Qy        374 VVPDHYCHYYPENVKPKPKLKECSMDPCPSSDCFKEIMPYDHFQPLPPWEHNPWTACSVS 433
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VVPDHYCHYYPENVKPKPKLKECSMDPCPSSDCFKEIMPYDHFQPLPPWEHNPWTACSVS 120

Qy        434 CGGGIQRRSFVCVEESMHGEILQVEEWKCMYAPKPKVMQTCNLFDCPKWIAMEWSQCTVT 493
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CGGGIQRRSFVCVEESMHGEILQVEEWKCMYAPKPKVMQTCNLFDCPKWIAMEWSQCTVT 180

Qy        494 CGRGLRYRVVLCINHRGEHVGGCNPQLKLHIKEECVIPIPCYKPKEKSPVEAKLPWLKQA 553
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CGRGLRYRVVLCINHRGEHVGGCNPQLKLHIKEECVIPIPCYKPKEKSPVEAKLPWLKQA 240

Qy        554 QELEETRIATEEPTFIPEPWSACSTTCGPGVQVREVKCRVLLTFTQTETELPEEECEGPK 613
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 QELEETRIATEEPTFIPEPWSACSTTCGPGVQVREVKCRVLLTFTQTETELPEEECEGPK 300

Qy        614 LPTERPCLLEACDESPASRELDIPLPEDSETTYDWEYAGFTPCTATCLGGHQEAIAVCLH 673
              |||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||
Db        301 LPTERPCLLEACDESPASRELDIPLPEDSETTYDWEYAGFTPCTATCVGGHQEAIAVCLH 360

Qy        674 IQTQQTVNDSLCDMVHRPPAMSQACNTEPCPPRWHVGSWGPCSATCGVGIQTRDVYCLHP 733
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 IQTQQTVNDSLCDMVHRPPAMSQACNTEPCPPRWHVGSWGPCSATCGVGIQTRDVYCLHP 420

Qy        734 GETPAPPEECRDEKPHALQACNQFDCPPGWHIEEWQQCSRTCGGGTQNRRVTCRQLLTDG 793
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GETPAPPEECRDEKPHALQACNQFDCPPGWHIEEWQQCSRTCGGGTQNRRVTCRQLLTDG 480

Qy        794 SFLNLSDELCQGPKASSHKSCARTDCPPHLAVGDWSKCSVSCGVGIQRRKQVCQRLAAKG 853
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 SFLNLSDELCQGPKASSHKSCARTDCPPHLAVGDWSKCSVSCGVGIQRRKQVCQRLAAKG 540

Qy        854 RRIPLSEMMCRDLPGFPLVRSCQMPECSKIKSEMKTKLGEQGPQILSVQRVYIQTREEKR 913
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 RRIPLSEMMCRDLPGFPLVRSCQMPECSKIKSEMKTKLGEQGPQILSVQRVYIQTREEKR 600

Qy        914 INLTIGSRAYLLPNTSVIIKCPVRRFQKSLIQWEKDGRCLQNSKRLGITKSGSLKIHGLA 973
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 INLTIGSRAYLLPNTSVIIKCPVRRFQKSLIQWEKDGRCLQNSKRLGITKSGSLKIHGLA 660

Qy        974 APDIGVYRCIAGSAQETVVLKLIGTDNRLIARPALREPMREYPGMDHSEANSLGVTWHKM 1033
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 APDIGVYRCIAGSAQETVVLKLIGTDNRLIARPALREPMREYPGMDHSEANSLGVTWHKM 720

Qy       1034 RQMWNNKNDLYLDDDHISNQPFLRALLGHCSNSAGSTNSWELKNKQFEAAVKQGAYSMDT 1093
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 RQMWNNKNDLYLDDDHISNQPFLRALLGHCSNSAGSTNSWELKNKQFEAAVKQGAYSMDT 780

Qy       1094 AQFDELIRNMSQLMETGEVSDDLASQLIYQLVAELAKAQPTHMQWRGIQEETPPAAQLRG 1153
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 AQFDELIRNMSQLMETGEVSDDLASQLIYQLVAELAKAQPTHMQWRGIQEETPPAAQLRG 840

Qy       1154 ETGSVSQSSHAKNSGKLTEKEKGPVLMRQSQPPSISNKTINSRIGNIVTITEKRTEVINI 1213
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ETGSVSQSSHAKNSGKLTEKEKGPVLMRQSQPPSISNKTINSRIGNIVTITEKRTEVINI 900

Qy       1214 LCDLITPSEATYTWTKDGTLLQPSVK 1239
              ||||||||||||||||||||||||||
Db        901 LCDLITPSEATYTWTKDGTLLQPSVK 926


RESULT 14
US-09-784-358-12
; Sequence 12, Application US/09784358
; GENERAL INFORMATION:
;  APPLICANT: Donoho, Gregory
;  APPLICANT:  Scoville, John
;  APPLICANT:  Turner, C. Alexander Jr.
;  APPLICANT:  Friedrich, Glenn
;  APPLICANT:  Zambrowicz, Brian
;  APPLICANT:  Sands, Arthur T.
;  TITLE OF INVENTION: NOVEL HUMAN THROMBOSPONDIN REPEAT PROTEINS AND
;  TITLE OF INVENTION:  POLYNUCLEOTIDES ENCODING THE SAME
;  FILE REFERENCE: LEX-0134-USA
;  CURRENT APPLICATION NUMBER: US/09/784,358
;  CURRENT FILING DATE:  2001-02-15
;  PRIOR APPLICATION NUMBER: US 60/183,282
;  PRIOR FILING DATE: 2000-02-17
;  NUMBER OF SEQ ID NOS: 17
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12
;   LENGTH: 845
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-784-358-12

  Query Match             50.2%;  Score 4699;  DB 21;  Length 845;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  830;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MASWTSPWWVLIGMVFMHSPLPQTTAEKSPGAYFLPEFALSPQGSFLEDTTGEQFLTYRY 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MASWTSPWWVLIGMVFMHSPLPQTTAEKSPGAYFLPEFALSPQGSFLEDTTGEQFLTYRY 60

Qy         61 DDQTSRNTRSDEDKDGNWDAWGDWSDCSRTCGGGASYSLRRCLTGRNCEGQNIRYKTCSN 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DDQTSRNTRSDEDKDGNWDAWGDWSDCSRTCGGGASYSLRRCLTGRNCEGQNIRYKTCSN 120

Qy        121 HDCPPDAEDFRAQQCSAYNDVQYQGHYYEWLPRYNDPAAPCALKCHAQGQNLVVELAPKV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 HDCPPDAEDFRAQQCSAYNDVQYQGHYYEWLPRYNDPAAPCALKCHAQGQNLVVELAPKV 180

Qy        181 LDGTRCNTDSLDMCISGICQAVGCDRQLGSNAKPDNCGVCAGDGSTCRLVRGQSKSHVSP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 LDGTRCNTDSLDMCISGICQAVGCDRQLGSNAKPDNCGVCAGDGSTCRLVRGQSKSHVSP 240

Qy        241 EKREENVIAVPLGSRSVRITVKGPAHLFIESKTLQGSKGEHSFNSPGVFVVENTTVEFQR 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 EKREENVIAVPLGSRSVRITVKGPAHLFIESKTLQGSKGEHSFNSPGVFVVENTTVEFQR 300

Qy        301 GSERQTFKIPGPLMADFIFKTRYTAAKDSVVQFFFYQPISHQWRQTDFFPCTVTCGGGYQ 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GSERQTFKIPGPLMADFIFKTRYTAAKDSVVQFFFYQPISHQWRQTDFFPCTVTCGGGYQ 360

Qy        361 LNSAECVDIRLKRVVPDHYCHYYPENVKPKPKLKECSMDPCPSSDCFKEIMPYDHFQPLP 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LNSAECVDIRLKRVVPDHYCHYYPENVKPKPKLKECSMDPCPSSDCFKEIMPYDHFQPLP 420

Qy        421 RWEHNPWTACSVSCGGGIQRRSFVCVEESMHGEILQVEEWKCMYAPKPKVMQTCNLFDCP 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 RWEHNPWTACSVSCGGGIQRRSFVCVEESMHGEILQVEEWKCMYAPKPKVMQTCNLFDCP 480
```

```
Qy        481 KWIAMEWSQCTVTCGRGLRYRVVLCINHRGEHVGGCNPQLKLHIKEECVIPIPCYKPKEK 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 KWIAMEWSQCTVTCGRGLRYRVVLCINHRGEHVGGCNPQLKLHIKEECVIPIPCYKPKEK 540

Qy        541 SPVEAKLPWLKQAQELEETRIATEEPTFIPEPWSACSTTCGPGVQVREVKCRVLLTFTQT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 SPVEAKLPWLKQAQELEETRIATEEPTFIPEPWSACSTTCGPGVQVREVKCRVLLTFTQT 600

Qy        601 ETELPEEECEGPKLPTERPCLLEACDESPASRELDIPLPEDSETTYDWEYAGFTPCTATC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ETELPEEECEGPKLPTERPCLLEACDESPASRELDIPLPEDSETTYDWEYAGFTPCTATC 660

Qy        661 LGGHQEAIAVCLHIQTQQTVNDSLCDMVHRPPAMSQACNTEPCPPRWHVGSWGPCSATCG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 LGGHQEAIAVCLHIQTQQTVNDSLCDMVHRPPAMSQACNTEPCPPRWHVGSWGPCSATCG 720

Qy        721 VGIQTRDVYCLHPGETPAPPEECRDEKPHALQACNQFDCPPGWHIEEWQQCSRTCGGGTQ 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 VGIQTRDVYCLHPGETPAPPEECRDEKPHALQACNQFDCPPGWHIEEWQQCSRTCGGGTQ 780

Qy        781 NRRVTCRQLLTDGSFLNLSDELCQGPKASSHKSCARTDCPPHLAVGDWSK 830
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 NRRVTCRQLLTDGSFLNLSDELCQGPKASSHKSCARTDCPPHLAVGDWSK 830


RESULT 15
US-09-784-358-14
; Sequence 14, Application US/09784358
; GENERAL INFORMATION:
;  APPLICANT: Donoho, Gregory
;  APPLICANT:  Scoville, John
;  APPLICANT:  Turner, C. Alexander Jr.
;  APPLICANT:  Friedrich, Glenn
;  APPLICANT:  Zambrowicz, Brian
;  APPLICANT:  Sands, Arthur T.
;  TITLE OF INVENTION:  NOVEL HUMAN THROMBOSPONDIN REPEAT PROTEINS AND
;  TITLE OF INVENTION:  POLYNUCLEOTIDES ENCODING THE SAME
;  FILE REFERENCE: LEX-0134-USA
;  CURRENT APPLICATION NUMBER: US/09/784,358
;  CURRENT FILING DATE:  2001-02-15
;  PRIOR APPLICATION NUMBER: US 60/183,282
;  PRIOR FILING DATE: 2000-02-17
;  NUMBER OF SEQ ID NOS: 17
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14
;   LENGTH: 771
;   TYPE: PRT
;   ORGANISM: homo sapiens
US-09-784-358-14

  Query Match             45.2%;  Score 4225;  DB 21;  Length 771;
  Best Local Similarity   91.1%;  Pred. No. 9.4e-298;
  Matches  756;  Conservative    0;  Mismatches    0;  Indels   74;  Gaps    1;

Qy          1 MASWTSPWWVLIGMVFMHSPLPQTTAEKSPGAYFLPEFALSPQGSFLEDTTGEQFLTYRY 60
              ||||||||||||||||||||||||||||||
Db          1 MASWTSPWWVLIGMVFMHSPLPQTTAEKSP------------------------------ 30

Qy         61 DDQTSRNTRSDEDKDGNWDAWGDWSDCSRTCGGGASYSLRRCLTGRNCEGQNTRYKTCSN 120
                                                          ||||||||||||||||
Db         31 --------------------------------------------GRNCEGQNTRYKTCSN 46

Qy        121 HDCPPDAEDFRAQQCSAYNDVQYQGHYYEWLPRYNDPAAPCALKCHAQGQNLVVELAPKV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         47 HDCPPDAEDFRAQQCSAYNDVQYQGHYYEWLPRYNDPAAPCALKCHAQGQNLVVELAPKV 106

Qy        181 LDGTRCNTDSLDMCISGICQAVGCDRQLGSNAKEDNCGVCAGDGSTCRLVRGQSKSHVSP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 LDGTRCNTDSLDMCISGICQAVGCDRQLGSNAKEDNCGVCAGDGSTCRLVRGQSKSHVSP 166

Qy        241 EKREENVIAVPLGSRSVRITVKGPAHLFIESKTLQGSKGEHSFNSPGVFVVENTTVEFQR 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        167 EKREENVIAVPLGSRSVRITVKGPAHLFIESKTLQGSKGEHSFNSPGVFVVENTTVEFQR 226

Qy        301 GSERQTFKIPGPLMADFIFKTRYTAAKDSVVQFFFYQPISHQWRQTDFFPCTVTCGGGYQ 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 GSERQTFKIPGPLMADFIFKTRYTAAKDSVVQFFFYQPISHQWRQTDFFPCTVTCGGGYQ 286

Qy        361 LNSAECVDIRLKRVVPDHYCHYYPENVKPKPKLKECSMDPCPSSDGFKEIMPYDHFQPLP 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 LNSAECVDIRLKRVVPDHYCHYYPENVKPKPKLKECSMDPCPSSDGFKEIMPYDHFQPLP 346

Qy        421 RWEHNPWTACSVSCGGGIQRRSFVCVEESMHGEILQVEEWKCMYAPKPKVMQTCNLFDCP 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        347 RWEHNPWTACSVSCGGGIQRRSFVCVEESMHGEILQVEEWKCMYAPKPKVMQTCNLFDCP 406

Qy        481 KWIAMEWSQCTVTCGRGLRYRVVLCINHRGEHVGGCNPQLKLHIKEECVIPIPCYKPKEK 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        407 KWIAMEWSQCTVTCGRGLRYRVVLCINHRGEHVGGCNPQLKLHIKEECVIPIPCYKPKEK 466

Qy        541 SPVEAKLPWLKQAQELEETRIATEEPTFIPEPWSACSTTCGPGVQVREVKCRVLLTFTQT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        467 SPVEAKLPWLKQAQELEETRIATEEPTFIPEPWSACSTTCGPGVQVREVKCRVLLTFTQT 526

Qy        601 ETELPEEECEGPKLPTERPCLLEACDESPASRELDIPLPEDSETTYDWEYAGFTPCTATC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        527 ETELPEEECEGPKLPTERPCLLEACDESPASRELDIPLPEDSETTYDWEYAGFTPCTATC 586

Qy        661 LGGHQEAIAVCLHIQTQQTVNDSLCDMVHRPPAMSQACNTEPCPPRWHVGSWGPCSATCG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        587 LGGHQEAIAVCLHIQTQQTVNDSLCDMVHRPPAMSQACNTEPCPPRWHVGSWGPCSATCG 646

Qy        721 VGIQTRDVYCLHPGETPAPPEECRDEKPHALQACNQFDCPPGWHIEEWQQCSRTCGGGTQ 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        647 VGIQTRDVYCLHPGETPAPPEECRDEKPHALQACNQFDCPPGWHIEEWQQCSRTCGGGTQ 706

Qy        781 NRRVTCRQLLTDGSFLNLSDELCQGPKASSHKSCARTDCPPHLAVGDWSK 830
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        707 NRRVTCRQLLTDGSFLNLSDELCQGPKASSHKSCARTDCPPHLAVGDWSK 756


Search completed: March 22, 2003, 15:36:29
Job time : 231 secs
```

```
                     GenCore version 5.1.4_p5_4578
              Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:          March 22, 2003, 02:30:08 ; Search time 9741 Seconds
                                             (without alignments)
                                             15165.367 Million cell updates/sec

Title:           US-09-784-358-1
Perfect score:   5076
Sequence:        1 atggcttcctggccaaggcc..........attcatgtcaaggggatataa 5076

Scoring table:   IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0

Searched:        2054640 seqs, 14155402878 residues

Total number of hits satisfying chosen parameters:       4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       GenEmbl:*
                 1:  gb_ba:*
                 2:  gb_htg:*
                 3:  gb_in:*
                 4:  gb_om:*
                 5:  gb_ov:*
                 6:  gb_pat:*
                 7:  gb_ph:*
                 8:  gb_pl:*
                 9:  gb_pr:*
                 10:  gb_ro:*
                 11:  gb_sts:*
                 12:  gb_sy:*
                 13:  gb_un:*
                 14:  gb_vi:*
                 15:  em_ba:*
                 16:  em_fun:*
                 17:  em_hum:*
                 18:  em_in:*
                 19:  em_mu:*
                 20:  em_om:*
                 21:  em_or:*
                 22:  em_ov:*
                 23:  em_pat:*
                 24:  em_ph:*
                 25:  em_pl:*
                 26:  em_ro:*
                 27:  em_sts:*
                 28:  em_un:*
                 29:  em_vi:*
                 30:  em_htg_hum:*
                 31:  em_htg_inv:*
                 32:  em_htg_other:*
                 33:  em_htg_mus:*
                 34:  em_htg_pln:*
                 35:  em_htg_rod:*
                 36:  em_htg_mam:*
                 37:  em_htg_vrt:*
                 38:  em_sy:*
                 39:  em_htgo_hum:*
                 40:  em_htgo_mus:*
                 41:  em_htgo_other:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                                  SUMMARIES
                     %
Result             Query
   No.    Score    Match  Length DB  ID                         Description
------------------------------------------------------------------------------
     1     5076    100.0    5076  6  AX224809                   AX224809 Sequence
     2   5069.6     99.9    7260  6  AX235356                   AX235356 Sequence
     3     4766     93.9    4854  6  AX224823                   AX224823 Sequence
     4   3862.2     76.1    6303  6  AX235410                   AX235410 Sequence
     5     3859     76.0    6294  6  AX235354                   AX235354 Sequence
     6     3755     74.0    8578  6  AX224825                   AX224825 Sequence
     7   3072.4     60.5    5107  6  AX235377                   AX235377 Sequence
     8   3072.4     60.5    5107  6  AX235379                   AX235379 Sequence
     9   3072.4     60.5    5107  6  AX235413                   AX235413 Sequence
    10   3072.4     60.5    5107  9  AB033059                   AB033059 Homo sapi
    11   3070.8     60.5    5107  6  AX235411                   AX235411 Sequence
    12   2862.8     56.4    2912  6  AX149471                   AX149471 Sequence
    13     2490     49.1    2538  6  AX224819                   AX224819 Sequence
    14   2295.8     45.2    2523  9  AF237652                   AF237652 Homo sapi
    15     2180     42.9    2316  6  AX224821                   AX224821 Sequence
    16   2120.6     41.8    4614  9  AB072736                   AB072736 Macaca fa
    17     2119     41.7    2175  6  AX224815                   AX224815 Sequence
    18     1809     35.6    1953  6  AX224817                   AX224817 Sequence
    19     1327     26.1    1341  6  AX224811                   AX224811 Sequence
c   20   1070.4     21.1  164879  9  AC022684                   AC022684 Homo sapi
c   21   1070.4     21.1  181327  9  AC027807                   AC027807 Homo sapi
c   22   1034.4     20.4  183976  2  AC073419                   AC073419 Homo sapi
    23     1017     20.0    1119  6  AX224813                   AX224813 Sequence
    24    803.4     15.8    2468  9  AK092324                   AK092324 Homo sapi
    25      738     14.5    1869  6  AX358900                   AX358900 Sequence
    26      738     14.5    1869  6  AX362393                   AX362393 Sequence
    27      738     14.5    1869  6  AX376368                   AX376368 Sequence
    28      738     14.5    1869  6  AX454558                   AX454558 Sequence
    29      738     14.5    1869  6  AX491036                   AX491036 Sequence
    30    733.2     14.4    1803  9  AF176313                   AF176313 Homo sapi
    31      536     10.6    1616  9  AF251058                   AF251058 Homo sapi
    32    530.6     10.5    2383  9  BC030262                   BC030262 Homo sapi
c   33    504.6      9.9  196726  2  AC126723                   AC126723 Rattus no
c   34    240.4      4.7  138028  2  AC103743                   AC103743 Homo sapi
c   35    240.4      4.7  212659  9  AC116157                   AC116157 Homo sapi
    36      219      4.3  165655  2  AC022605                   AC022605 Homo sapi
    37    191.6      3.8  183976  2  AC073419                   AC073419 Homo sapi
c   38    182.8      3.6  170140  2  AC110192                   AC110192 Mus muscu
    39    180.6      3.6   64303  2  AC103754                   AC103754 Homo sapi
    40      173      3.4  168497  2  AC092410                   AC092410 Bos tauru
    41    167.8      3.3  196726  2  AC126723                   AC126723 Rattus no
    42    154.4      3.0     459  5  AF469478                   AF469478 Ajaja aja
    43    145.4      2.9  213216  9  HS990M15                   AL449063 Homo sapi
c   44    142.6      2.8   67308  2  AC110887                   AC110887 Mus muscu
c   45    142.2      2.8  142319  2  AC118449                   AC118449 Rattus no


                                  ALIGNMENTS


RESULT 1
AX224809
LOCUS       AX224809                5076 bp    DNA     linear   PAT 14-SEP-2001
DEFINITION  Sequence 1 from Patent WO0164011.
ACCESSION   AX224809
VERSION     AX224809.1  GI:15554609
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 5076)
  AUTHORS   Donoho,G., Scoville,J., Turner,C.A., Friedrich,G., Zambrowicz,B.
            and Sands,A.T.
  TITLE     Novel human thrombospondin repeat proteins and polynucleotides
```

```
            encoding the same
  JOURNAL   Patent: WO 0161011-A 1 23-AUG-2001;
            Lexicon Genetics Incorporated (US)
FEATURES             Location/Qualifiers
     source          1..5076
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT     1384 a   1227 c   1338 g   1127 t
ORIGIN

  Query Match             100.0%;  Score 5076;  DB 6;  Length 5076;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 5076;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGGCTTCCTGGACGAGCCCCTGGTGGGTGCTGATAGGCATGGTCTTCATGCACTCTCCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGCTTCCTGGACGAGCCCCTGGTGGGTGCTGATAGGCATGGTCTTCATGCACTCTCCC 60

Qy         61 CTCCCGCAGACCACAGCTGAGAAATCTCCTGGAGCCTATTTCCTTCCCGAGTTTGCACTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTCCCGCAGACCACAGCTGAGAAATCTCCTGGAGCCTATTTCCTTCCCGAGTTTGCACTT 120

Qy        121 TCTCCTCAGGGAAGTTTTCTGGAAGACACAACAGGGGAGCAGTTCCTCACTTATCGCTAT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCTCCTCAGGGAAGTTTTCTGGAAGACACAACAGGGGAGCAGTTCCTCACTTATCGCTAT 180

Qy        181 GATGACCAGACCTCAAGAAACACTCGTTCAGATGAAGACAAAGATGGCAACTGGGATGCT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GATGACCAGACCTCAAGAAACACTCGTTCAGATGAAGACAAAGATGGCAACTGGGATGCT 240

Qy        241 TGGGGCGACTGGAGTGACTGCTCCCGGACCTGTGGGGGAGGAGCATCATATTCTCTGCGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGGGGCGACTGGAGTGACTGCTCCCGGACCTGTGGGGGAGGAGCATCATATTCTCTGCGG 300

Qy        301 AGATGTTTGACTGGAAGGAATTGTGAAGGGCAGAACATTCGGTACAAGACATGCAGCAAT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGATGTTTGACTGGAAGGAATTGTGAAGGGCAGAACATTCGGTACAAGACATGCAGCAAT 360

Qy        361 CATGACTGCCCTCCAGATGCAGAAGATTTCAGAGCCCAGCAGTGCTCAGCCTACAATGAT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CATGACTGCCCTCCAGATGCAGAAGATTTCAGAGCCCAGCAGTGCTCAGCCTACAATGAT 420

Qy        421 GTCCAGTATCAGGGGCATTACTATGAATGGCTTCCACGATATAATGATCCTGCTGCCCCG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GTCCAGTATCAGGGGCATTACTATGAATGGCTTCCACGATATAATGATCCTGCTGCCCCG 480

Qy        481 TGTGCACTCAAGTGTCATGCACAAGGACAAAACTTGGTGGTGGAGCTGGCACCTAAGGTA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TGTGCACTCAAGTGTCATGCACAAGGACAAAACTTGGTGGTGGAGCTGGCACCTAAGGTA 540

Qy        541 CTGGATGGAACTCGTTGCAACACGGACTCCTTGGACATGTGTATCAGTGGCATCTGTCAG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTGGATGGAACTCGTTGCAACACGGACTCCTTGGACATGTGTATCAGTGGCATCTGTCAG 600

Qy        601 GCAGTGGGCTGCGATCGGCAACTGGGAAGCAATGCCAAGGAGGACAACTGTGGAGTCTGT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GCAGTGGGCTGCGATCGGCAACTGGGAAGCAATGCCAAGGAGGACAACTGTGGAGTCTGT 660

Qy        661 GCCGGCGATGGCTCCACCTGCAGGCTTGTACGGGGACAATCAAAGTCACACGTTTCTCCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GCCGGCGATGGCTCCACCTGCAGGCTTGTACGGGGACAATCAAAGTCACACGTTTCTCCT 720

Qy        721 GAAAAAAGAGAAGAAAATGTAATTGCTGTTCCTTTGGGAAGTCGAAGTGTGAGAATTACA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GAAAAAAGAGAAGAAAATGTAATTGCTGTTCCTTTGGGAAGTCGAAGTGTGAGAATTACA 780

Qy        781 GTGAAAGGACCTGCCCACCTCTTTATTGAATCAAAAACACTTCAAGGAAGCAAAGGAGAA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GTGAAAGGACCTGCCCACCTCTTTATTGAATCAAAAACACTTCAAGGAAGCAAAGGAGAA 840

Qy        841 CACAGCTTTAACAGCCCCGGCGTCTTTGTCGTAGAAAACACAACAGTGGAATTTCAGAGG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CACAGCTTTAACAGCCCCGGCGTCTTTGTCGTAGAAAACACAACAGTGGAATTTCAGAGG 900

Qy        901 GGCTCCGAGAGGCAAACTTTTAAGATTCCAGGACCTCTGATGGCTGATTTCATCTTCAAG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GGCTCCGAGAGGCAAACTTTTAAGATTCCAGGACCTCTGATGGCTGATTTCATCTTCAAG 960

Qy        961 ACCAGGTACACTGCAGCCAAAGACAGCGTGGTTCAGTTCTTCTTTTACCAGCCCATCAGT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 ACCAGGTACACTGCAGCCAAAGACAGCGTGGTTCAGTTCTTCTTTTACCAGCCCATCAGT 1020

Qy       1021 CATCAGTGGAGACAAACTGACTTCTTTCCCTGCACTGTGACGTGTGGAGGAGGTTATCAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CATCAGTGGAGACAAACTGACTTCTTTCCCTGCACTGTGACGTGTGGAGGAGGTTATCAG 1080

Qy       1081 CTCAATTCTGCTGAATGTGTGGATATCCGCTTGAAGAGGGTAGTTCCTGACCATTATTGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CTCAATTCTGCTGAATGTGTGGATATCCGCTTGAAGAGGGTAGTTCCTGACCATTATTGT 1140

Qy       1141 CACTACTACCCTGAAAATGTAAAACCAAAACCAAAACTGAAGGAATGCAGCATGGATCCC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CACTACTACCCTGAAAATGTAAAACCAAAACCAAAACTGAAGGAATGCAGCATGGATCCC 1200

Qy       1201 TGCCCATCAAGTGATGGATTTAAAGAGATAATGCCCTATGACCACTTCCAACCTCTTCCT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TGCCCATCAAGTGATGGATTTAAAGAGATAATGCCCTATGACCACTTCCAACCTCTTCCT 1260

Qy       1261 CGCTGGGAACATAATCCTTGGACTGCATGTTCCGTGTCCTGTGGAGGAGGGATTCAGAGA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CGCTGGGAACATAATCCTTGGACTGCATGTTCCGTGTCCTGTGGAGGAGGGATTCAGAGA 1320

Qy       1321 CGGAGCTTTGTGTGTGTAGAGGAATCCATGCATGGAGAGATATTGCAGGTGGAAGAATGG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CGGAGCTTTGTGTGTGTAGAGGAATCCATGCATGGAGAGATATTGCAGGTGGAAGAATGG 1380

Qy       1381 AAGTGCATGTACGCACCCAAACCCAAGGTTATGCAAACTTGTAATCTGTTTGATTGCCCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 AAGTGCATGTACGCACCCAAACCCAAGGTTATGCAAACTTGTAATCTGTTTGATTGCCCC 1440

Qy       1441 AAGTGGATTGCCATGGAGTGGTCTCAGTGCACAGTGACTTGTGGCCGAGGGTTACGGTAC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 AAGTGGATTGCCATGGAGTGGTCTCAGTGCACAGTGACTTGTGGCCGAGGGTTACGGTAC 1500

Qy       1501 CGGGTTGTTCTGTGTATTAACCACCGCGGAGAGCATGTTGGGGGCTGCAATCCACAACTG 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 CGGGTTGTTCTGTGTATTAACCACCGCGGAGAGCATGTTGGGGGCTGCAATCCACAACTG 1560

Qy       1561 AAGTTACACATCAAAGAAGAATGTGTCATTCCCATCCCGTGTTATAAACCAAAAGAAAAA 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 AAGTTACACATCAAAGAAGAATGTGTCATTCCCATCCCGTGTTATAAACCAAAAGAAAAA 1620

Qy       1621 AGTCCAGTGGAAGCAAAATTGCCTTGGCTGAAACAAGCACAAGAACTAGAAGAGACCAGA 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 AGTCCAGTGGAAGCAAAATTGCCTTGGCTGAAACAAGCACAAGAACTAGAAGAGACCAGA 1680

Qy       1681 ATAGCAACAGAAGAACCAACGTTCATTCCAGAACCCTGGTCAGCCTGCAGTACCACGTGT 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 ATAGCAACAGAAGAACCAACGTTCATTCCAGAACCCTGGTCAGCCTGCAGTACCACGTGT 1740

Qy       1741 GGGCCAGGTGTGCAGGTCCGCGAGGTGAAGTGCCGTGTGCTCCTCACATTCACGCAGACT 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 GGGCCAGGTGTGCAGGTCCGCGAGGTGAAGTGCCGTGTGCTCCTCACATTCACGCAGACT 1800

Qy       1801 GAGACTGAGCTGCCCGAGGAAGAGTGTGAAGGCCCCAAGCTGCCCACCGAACGGCCCTGC 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 GAGACTGAGCTGCCCGAGGAAGAGTGTGAAGGCCCCAAGCTGCCCACCGAACGGCCCTGC 1860

Qy       1861 CTCCTGGAAGCATGTGATGAGAGCCCGGCCTCCCGAGAGCTAGACATCCCTCTCCCTGAG 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 CTCCTGGAAGCATGTGATGAGAGCCCGGCCTCCCGAGAGCTAGACATCCCTCTCCCTGAG 1920

Qy       1921 GACAGTGAGACGACTTACGACTGGGAGTACGCTGGGTTCACCCCTTGCACAGCAACATGC 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 GACAGTGAGACGACTTACGACTGGGAGTACGCTGGGTTCACCCCTTGCACAGCAACATGC 1980
```

```
Qy       1981 TTGGGAGGCCATCAAGAAGCCATAGCAGTGTGCTTACATATCCAGACCCAGCAGACAGTC 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 TTGGGAGGCCATCAAGAAGCCATAGCAGTGTGCTTACATATCCAGACCCAGCAGACAGTC 2040

Qy       2041 AATGACAGCTTGTGTGATATGGTCCACCGTGCCTCAGCCATGAGCCAGGCCTGTAACACA 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 AATGACAGCTTGTGTGATATGGTCCACCGTGCCTCAGCCATGAGCCAGGCCTGTAACACA 2100

Qy       2101 GAGCCCTGTCCCCCCAGGTGGCATGTGGGCTCTTGGGGGCCCTGCTCAGCTACCTGTCCA 2160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 GAGCCCTGTCCCCCCAGGTGGCATGTGGGCTCTTGGGGGCCCTGCTCAGCTACCTGTCCA 2160

Qy       2161 GTTGGAATTCAGACCCGAGATGTGTACTGCCTGCACCCAGGGGAGACCCCTGCCCCTCCT 2220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 GTTGGAATTCAGACCCGAGATGTGTACTGCCTGCACCCAGGGGAGACCCCTGCCCCTCCT 2220

Qy       2221 GAGGAGTGCCGAGATGAAAAGCCCCATGCTTTACAAGCATGCAATCAGTTTGACTGCCCT 2280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 GAGGAGTGCCGAGATGAAAAGCCCCATGCTTTACAAGCATGCAATCAGTTTGACTGCCCT 2280

Qy       2281 CCTGGCTGGCACATTGAAGAATGGCAGCAGTGTTCCAGGACTTGTGGCGGGGGAACTCAG 2340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 CCTGGCTGGCACATTGAAGAATGGCAGCAGTGTTCCAGGACTTGTGGCGGGGGAACTCAG 2340

Qy       2341 AACAGAAGAGTCACCTTGTGGCAGCTGCTAACGGATGGCAGCTTTTGAATCTCTCCAGAT 2400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 AACAGAAGAGTCACCTTGTGGCAGCTGCTAACGGATGGCAGCTTTTGAATCTCTCCAGAT 2400

Qy       2401 GAATTGTGCCAAGGACCCAAGGCATCGTCTCACAAGTCCTGTGCCAGGACAGACTGTCCT 2460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 GAATTGTGCCAAGGACCCAAGGCATCGTCTCACAAGTCCTGTGCCAGGACAGACTGTCCT 2460

Qy       2461 CCACATTTAGCTGTGGGAGACTGGTCGAAGTGTTCTGTCAGTTGTGGTGTGGGAATCCAG 2520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 CCACATTTAGCTGTGGGAGACTGGTCGAAGTGTTCTGTCAGTTGTGGTGTGGGAATCCAG 2520

Qy       2521 AGAAGAAAGCAGGTTGTGTCAAAGGCTGGCAGCCAAAGGTCGGCGCATCCCCTCAGTGAG 2580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 AGAAGAAAGCAGGTTGTGTCAAAGGCTGGCAGCCAAAGGTCGGCGCATCCCCTCAGTGAG 2580

Qy       2581 ATGATGTGCCAGGGATCTACCAGGGTCCCTCTTGTAAGATCTTGCCAGATGCCTGAGTGC 2640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 ATGATGTGCCAGGGATCTACCAGGGTCCCTCTTGTAAGATCTTGCCAGATGCCTGAGTGC 2640

Qy       2641 AGTAAAATCAAATCAGAGATGAAGACAAAACTTGGTGAGCAGGGTCCCCAGATCCTCAGT 2700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 AGTAAAATCAAATCAGAGATGAAGACAAAACTTGGTGAGCAGGGTCCCCAGATCCTCAGT 2700

Qy       2701 GTCCAGAGAGTCTACATTCAGACAAGGGAAGAGAAGCGTATTAACCTGACCATTGGTAGC 2760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 GTCCAGAGAGTCTACATTCAGACAAGGGAAGAGAAGCGTATTAACCTGACCATTGGTAGC 2760

Qy       2761 AGAGCCTATTTGCTGCCCAACACATCCGTGATTATTAAGTGCCCCGTGCGACGATTCCAG 2820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 AGAGCCTATTTGCTGCCCAACACATCCGTGATTATTAAGTGCCCCGTGCGACGATTCCAG 2820

Qy       2821 AAATCTCTGATCCAGTGGGAGAAGGATGGCCGTTGCCTGCAGAACTCCAAACGGCTTGGC 2880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 AAATCTCTGATCCAGTGGGAGAAGGATGGCCGTTGCCTGCAGAACTCCAAACGGCTTGGC 2880

Qy       2881 ATCACCAAGTCAGGCTCACTAAAAATCCACGGTCTTGCTGCCCCCGACATCGGCGTGTAC 2940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 ATCACCAAGTCAGGCTCACTAAAAATCCACGGTCTTGCTGCCCCCGACATCGGCGTGTAC 2940

Qy       2941 CGGTGCATTGCAGGCTCTGCACAGGAAACAGTTGTGCTCAAGCTCATTGGTACTGACAAC 3000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 CGGTGCATTGCAGGCTCTGCACAGGAAACAGTTGTGCTCAAGCTCATTGGTACTGACAAC 3000

Qy       3001 CGGCTCATCGCACGCCCAGCCCCAGGGAGCCTATGAGGGAATATCCTGGGATGGAGCCAC 3060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 CGGCTCATCGCACGCCCAGCCCCAGGGAGCCTATGAGGGAATATCCTGGGATGGAGCCAC 3060

Qy       3061 AGCGAAGCCAATAGTTTGGGAGTCACATGGCACAAAATGAGGCAAATGTGGAATAACAAA 3120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 AGCGAAGCCAATAGTTTGGGAGTCACATGGCACAAAATGAGGCAAATGTGGAATAACAAA 3120

Qy       3121 AATGACTTTATCTGGATGATGACCACATTAGTAACCAGCCTTTCTTGACACCTCTTGTTA 3180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3121 AATGACTTTATCTGGATGATGACCACATTAGTAACCAGCCTTTCTTGACACCTCTTGTTA 3180

Qy       3181 GGCCACTGCAGCAATTCTGCAGGAAGCACCAACTCCTGGGAGTTGAAGAATAAGCAGTTT 3240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3181 GGCCACTGCAGCAATTCTGCAGGAAGCACCAACTCCTGGGAGTTGAAGAATAAGCAGTTT 3240

Qy       3241 GAAGCAGCAGTTAAACAAGGAGCATATACCATGGATACATCCCAGTTTGATCAGCTGATA 3300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3241 GAAGCAGCAGTTAAACAAGGAGCATATACCATGGATACATCCCAGTTTGATCAGCTGATA 3300

Qy       3301 AGAAACATGAGTCAGCTCATGGAAACCGGAGAGGTCAGCGATGATCTTGCGTCCCAGCTG 3360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3301 AGAAACATGAGTCAGCTCATGGAAACCGGAGAGGTCAGCGATGATCTTGCGTCCCAGCTG 3360

Qy       3361 ATATATCAGCTGGTGGCCGAATTAGCCAAGGCACAGCCAACACACATGCAGTGCGGGGGC 3420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3361 ATATATCAGCTGGTGGCCGAATTAGCCAAGGCACAGCCAACACACATGCAGTGCGGGGGC 3420

Qy       3421 ATCCAGGAAGAGACACCTTCCTGCTGCTCAGCTCAGAGGGGAAACAGGGAGTGTGTCCAA 3480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3421 ATCCAGGAAGAGACACCTTCCTGCTGCTCAGCTCAGAGGGGAAACAGGGAGTGTGTCCAA 3480

Qy       3481 AGCTCGGCATGCAAAAAACTCAGGCAAGCTGACATTCAAGCCGAAAGGACTGTTCTCATG 3540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3481 AGCTCGGCATGCAAAAAACTCAGGCAAGCTGACATTCAAGCCGAAAGGACTGTTCTCATG 3540

Qy       3541 AGGCAAAGCCAACCTCCCTCAATTTCATTTAATAAAACAATAAATTCCAGGATTGGAAAT 3600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3541 AGGCAAAGCCAACCTCCCTCAATTTCATTTAATAAAACAATAAATTCCAGGATTGGAAAT 3600

Qy       3601 ACAGTATACATTACAAAAAGGACAGAGGTCATCAATATACTGTGTGACCTTATTACCCCC 3660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3601 ACAGTATACATTACAAAAAGGACAGAGGTCATCAATATACTGTGTGACCTTATTACCCCC 3660

Qy       3661 AGTGAGGCCACATATACATGGACCAAGGATGGAACCTTGTTACAGCCCTCAGTAAAAATA 3720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3661 AGTGAGGCCACATATACATGGACCAAGGATGGAACCTTGTTACAGCCCTCAGTAAAAATA 3720

Qy       3721 ATTTTGGATGGAACTGGGAAGATACAGATACAGAATCCTACAAGGAAACAACAAGGCATA 3780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3721 ATTTTGGATGGAACTGGGAAGATACAGATACAGAATCCTACAAGGAAACAACAAGGCATA 3780

Qy       3781 TATGAATGTTCTTTAGCTAATCATCTTGGTTCAGATGTGGAAAGTTCTTCTGCCCTGTAT 3840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3781 TATGAATGTTCTTTAGCTAATCATCTTGGTTCAGATGTGGAAAGTTCTTCTGCCCTGTAT 3840

Qy       3841 GCAGAGGCACCTGTTCATCTTGTTCTGTTGAAAGAAATATCACAAACAGAGCACAACCAT 3900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3841 GCAGAGGCACCTGTTCATCTTGTTCTGTTGAAAGAAATATCACAAACAGAGCACAACCAT 3900

Qy       3901 CTGTCTGTTGTGGTTGGAGGCATCGTGGAGGCAGCCCTTGGAGCAAACGTGACAATCCGA 3960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3901 CTGTCTGTTGTGGTTGGAGGCATCGTGGAGGCAGCCCTTGGAGCAAACGTGACAATCCGA 3960

Qy       3961 TGTCCTGTAAAAGGTGTCCCTCAGCCTAATATAACTTGGTTGAAGAGAGGAGGATCTCTG 4020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3961 TGTCCTGTAAAAGGTGTCCCTCAGCCTAATATAACTTGGTTGAAGAGAGGAGGATCTCTG 4020

Qy       4021 AGTGGCAATGTTTCCTTGCTTCTCAATGGATCCCTGTTGTTGCAGAATGTTTCCCTTGAA 4080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4021 AGTGGCAATGTTTCCTTGCTTCTCAATGGATCCCTGTTGTTGCAGAATGTTTCCCTTGAA 4080

Qy       4081 AATGAAGGAACCTACGTCTGCATAGCCACCAATGCTCTTGGAAAGGCAGTGGCAACATCT 4140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4081 AATGAAGGAACCTACGTCTGCATAGCCACCAATGCTCTTGGAAAGGCAGTGGCAACATCT 4140

Qy       4141 GTACTCCACTTGCTGGAAGGAAGATGGCCAGAGATAGAATCGTATTCTGCTGAAGGACAT 4200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4141 GTACTCCACTTGCTGGAACGAAGATGGCCAGAGAGTAGAATCGTATTTCTGCAAGGACAT 4200

Qy       4201 AAAAAGTACATTCTCCAGGCAACCAACACTAGAACCAACAGCAATGACCCAACAGGAGAA 4260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4201 AAAAAGTACATTCTCCAGGCAACCAACACTAGAACCAACAGCAATGACCCAACAGGAGAA 4260

Qy       4261 CCCCCGCCTCAAGAGCCTTTTTGGGAGCCTGGTAACTGGTCACATTGTTCTGCCACCTGT 4320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4261 CCCCCGCCTCAAGAGCCTTTTTGGGAGCCTGGTAACTGGTCACATTGTTCTGCCACCTGT 4320

Qy       4321 GGTCATTTGGGAGCCCGCATTCAGAGACCCCAGTGTGTGATGGCCAATGGGCAGGAAGTG 4380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4321 GGTCATTTGGGAGCCCGCATTCAGAGACCCCAGTGTGTGATGGCCAATGGGCAGGAAGTG 4380

Qy       4381 AGTGAGGCCCTGTGTGATCACCTCCAGAAGCCACTGGCTGGGTTTGAGCCCTGTAACATC 4440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4381 AGTGAGGCCCTGTGTGATCACCTCCAGAAGCCACTGGCTGGGTTTGAGCCCTGTAACATC 4440

Qy       4441 CGGGACTGCCCAGCGAGGTGGTTCACAAGTGTGTGGTCACAGTGCTCTGTGTCTTGCGCT 4500
              |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db       4441 CGGGACTGCCCAGCGAGGTGGTTCACAAGTGTGTGGTCACAGTGCTCTGTGTCTTGCGGT 4500

Qy       4501 GAAGGATACCACAGTCGGCAGGTGACGTGCAAGCGGCACAAAACCAATGGAACTGTGCAC 4560
              |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db       4501 GAAGGATACCACAGTCGGCAGGTGACGTGCAAGCGGACAAAAGCCAATGGAACTGTGCAG 4560

Qy       4561 GTGGTGTCTCCAAGAGCATGTGCCCCTAAAGACCGGCCTCTGGGAAGAAAACCATGTTTT 4620
              |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db       4561 GTGGTGTCTCCAAGAGCATGTGCCCCTAAAGACCGGCCTCTGGGAAGAAAACCATGTTTT 4620

Qy       4621 GGTCATCCATGTGTTCAGTGGGAACCAGGGAACCGGTGTCCTGGACGTTGCATGGGCCGT 4680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4621 GGTCATCCATGTGTTCAGTGGGAACCAGGGAACCGGTGTCCTGGACGTTGCATGGGCCGT 4680

Qy       4681 GCTGTGAGGATGCAGCAGCGTCACACAGCTTGTCAACACAACAGCTCTGACTCCAACTGT 4740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4681 GCTGTGAGGATGCAGCAGCGTCACACAGCTTGTCAACACAACAGCTCTGACTCCAACTGT 4740

Qy       4741 GATGACAGAAAGAGACCCACCTTAAGAAGGAACTGCACATCAGGGGCCTGTGATGTGTGT 4800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4741 GATGACAGAAAGAGACCCACCTTAAGAAGGAACTGCACATCAGGGGCCTGTGATGTGTGT 4800

Qy       4801 TGGCACACAGGCCCTTGGAAGCCCTCTACAGCAGCCTGTGGCACGGCTTTCCAGTCTCCC 4860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4801 TGGCACACAGGCCCTTGGAAGCCCTCTACAGCAGCCTGTGGCACGGCTTTCCAGTCTCCC 4860

Qy       4861 AAAGTCGACTGTATCCACACAAGGAGTTGCAAACCTGTGGCCAAGACACACTGTGTACAC 4920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4861 AAAGTCGACTGTATCCACACAAGGAGTTGCAAACCTGTGGCCAAGACACACTGTGTACAG 4920

Qy       4921 AAAAAGAAACCAATTTCCTGGCGGCACTGTCTTGGGCCCTCCTGTGATAGAGACTGCACA 4980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4921 AAAAAGAAACCAATTTCCTGGCGGCACTGTCTTGGGCCCTCCTGTGATAGAGACTGCACA 4980

Qy       4981 GACACAACTCACTACTGTATGTTTGTAAAACATCTTAATTTGTGTTCTCTAGACCGCTAC 5040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4981 GACACAACTCACTACTGTATGTTTGTAAAACATCTTAATTTGTGTTCTCTAGACCGCTAC 5040

Qy       5041 AAACAAAGGTGCTGCCAGTCATGTCAAGAGGGATAA 5076
              ||||||||||||||||||||||||||||||||||||
Db       5041 AAACAAAGGTGCTGCCAGTCATGTCAAGAGGGATAA 5076


RESULT 2
AX235356
LOCUS       AX235356                7260 bp    DNA     linear   PAT 11-SEP-2001
DEFINITION  Sequence 3 from Patent WO0162928.
ACCESSION   AX235356
VERSION     AX235356.1  GI:15593889
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 7260)
  AUTHORS   Vernet,C.A., Fernandes,E., Shimkets,R.A., Macdougall,J. and
            Spaderna,S.K.
  TITLE     Polypeptides and nucleic acids encoding same
  JOURNAL   Patent: WO 0162928-A 3 30 AUG 2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1..7260
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             136..5211
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC69812.1"
                     /db_xref="GI:15593890"
                     /translation="MASWTSPWWVLIGMVFMHSPLPQTTAEKSPGAYFLPEFALSPQG
                     SFLEDTTGEQFLTYRYDDQTSRNTRSDEDKDGNWDAWGDWSDCSRTCGGGASYSLRRC
                     LTGRNCEGQNIRYKTCSNHDCPPDAEDFRAQQCSAYNDVQYQGHYYEWLPRYNDPAAP
                     CALKCHAQGQNLVVELAPKVLDGTRCNTDSLDMCISGICQAVGCDRQLGSNAKEDNCG
                     VCAGDGSTCRLVRGQSKSHVSPEKREENVIAVPLGSRSVRITVKGPAHLFIESKTLQG
                     SEGPHCPNSPIVSVVLN[illegible]LMA[illegible]AAHDSVVQFF
                     FYQ[illegible]VI[illegible]YQLNSAH[illegible]LKFVV[illegible]CHYYPENVKPKPK
                     LKECSMDPCPSSDGFKEIMPYDHFQPLPRWEHNPWTACSVSCGGGIQRRSFVCVEESM
                     HGEILQVEEWKCMYAPKPKVMQTCNLFDCPKWIAMEWSQCTVTCGRGLRYRVVLCINH
                     RGEHVGGCNPQLKLHIKEECVIPIPCYKPKEKSPVEAKLPWLKQAQELEETRIATEEP
                     TFIPEEWSACSTTCGPGVQVREVKCRVLLTFTQTETELPEEECEGPKLPTERPCLLEA
                     CDESPASRELDIPLPEDSETTYDWEYAGFTPCTATCVGGHQEATAVCLHIQTQQTVND
                     SLCDMVHRPPAMSQACNTEPCPPRWHVGSWGPCSATCGVGIQTRDVYCLHPGETPAPP
                     EECRDEKPHALQACNQFDCPPGWHIEEWQQCSRTCGGGTQNRRVTCRQLLTDGSFLNL
                     SDELCQGPKASSHKSCARTDCPPHLAVGDWSKCSVSCGVGIQRRKQVCQRLAAKGRRI
                     PLSEMMCRDLPGFPLVRSCQMPECSKIKSEMKTKLGEQGPQILSVQRVYIQTREEKRI
                     NLTI[illegible]THGL
                     AAPDIGVTRCIAGSAQETVVLKLIGTDNKLIARLALKEPMKEYIGMDGHSLANSLGVIW
                     HEMEQMWNNENDLYLLP[illegible]AAVKQGA
                     YSMDTAQFDELIRNMSQLMETGEVSDDLASQLIYQLVAELAKAQPTHMQWRGIQEE[illegible]
                     PAAQ[illegible]FNKTINSRIGNTVYI
                     TKRTEVINILCDLITPSFATYTWTKDGTLLQPSVKIILDGTGKIQIQNPTRKEQGIYE
                     CSVANHLGSDVESSSVLYAEAPVILSVERNITKPEHNHLSVVVGGIVEAALGANVTIR
                     CPVKGVPQPNITWLKRGGSLSGNVSLLFNGSLLLQNVSLENEGTYVCIATNALGKAVA
                     TSVFHLLERRWPESRIVFLQGHKKYILQATNIRTNSNDPTGEPPPQEPPWEPGNWSHC
                     SATCGHLGARIQRPQCVMANGQEVSEALCDHLQKPLAGFEPCNIRDCPARWFTSVWSQ
                     [illegible]GNR
                     CPGRCMGRAVRMQRHTACQHNSC[illegible]WKPCT
                     AACGPCGQSKKV[illegible]MF
                     VKHLNLCSLDRYKQRCCQSCQEG"
BASE COUNT     2022 a   1593 c   1811 g   1834 t
ORIGIN

  Query Match             99.9%;  Score 5069.6;  DB 6;  Length 7260;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 5072;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 ATGGCTTCCTGGACGAGCCCCTGGTGGGTGCTGATAGGGATGGTCTTCATGCACTCTCCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        136 ATGGCTTCCTGGACGAGCCCCTGGTGGGTGCTGATAGGGATGGTCTTCATGCACTCTCCC 195

Qy         61 CTCCCGGCAGACCACAGCTGAGAAATCTCCTGGAGCCTATTTCCTTCCCGAGTTTGCACTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        196 CTCCCGGCAGACCACAGCTGAGAAATCTCCTGGAGCCTATTTCCTTCCCGAGTTTGCACTT 255

Qy        121 TCTCCTCAGGGAAGTTTTCTGGAAGACACAACAGGGGAGCAGTTCCTCACTTATCGCTAT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        256 TCTCCTCAGGGAAGTTTTCTGGAAGACACAACAGGGGAGCAGTTCCTCACTTATCGCTAT 315

Qy        181 GATGACCAGACCTCAAGAAACACTCGTTCAGATGAAGACAAAGATGGCAACTGGGATGCT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        316 GATGACCAGACCTCAAGAAACACTCGTTCAGATGAAGACAAAGATGGCAACTGGGATGCT 375

Qy        241 TGGGGCGACTGGAGTGACTGCTCCCGGACCTGTGGGGGAGGAGCATCATATTCTCTGCGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        376 TGGGGCGACTGGAGTGACTGCTCCCGGACCTGTGGGGGAGGAGCATCATATTCTCTGCGG 435
```

```
Qy        301 AGATGTTTGACTGGAAGGAATTGTGAAGGGCAGAACATTCGGTACAAGACATGCAGCAAT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        436 AGATGTTTGACTGGAAGGAATTGTGAAGGGCAGAACATTCGGTACAAGACATGCAGCAAT 495

Qy        361 CATGACTGGCCTCCAGATCCAGAAGATTTCAGAGCCCAGCAGTGCTCAGCCTACAATGAT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        496 CATGACTGGCCTCCAGATCCAGAAGATTTCAGAGCCCAGCAGTGCTCAGCCTACAATGAT 555

Qy        421 GTCCAGTATCAGGGCATTACTATGAATGGCTTCCACGATATAATGATCCTGCTGCCCCG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        556 GTCCAGTATCAGGGCATTACTATGAATGGCTTCCACGATATAATGATCCTGCTGCCCCG 615

Qy        481 TGTGCACTCAAGTGTCATGCACAAGGACAAAACTTGGTGGTGGAGCTGGCACCTAAGGTA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        616 TGTGCACTCAAGTGTCATGCACAAGGACAAAACTTGGTGGTGGAGCTGGCACCTAAGGTA 675

Qy        541 CTGGATGGAACTGGTTGCAACACGGACTCCTTGGACATGTGTATCAGTGGCATCTGTCAG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        676 CTGGATGGAACTGGTTGCAACACGGACTCCTTGGACATGTGTATCAGTGGCATCTGTCAT 735

Qy        601 GCAGTGGGCTGCGATCGGCAACTGGGAAGCAATGCCAAGGAGGACAACTGTGGAGTCTGT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        736 GCAGTGGGCTGCGATCGGCAACTGGGAAGCAATGCCAAGGAGGACAACTGTGGAGTCTGT 795

Qy        661 GCCGGCGATGGCTCCACCTGCAGGCTTGTACGGGGACAATCAAAGTCACAGGTTTCTCCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        796 GCCGGCGATGGCTCCACCTGCAGGCTTGTACGGGGACAATCAAAGTCACAGGTTTCTCCT 855

Qy        721 GAAAAAAGAGAAGAAAATGTAATTGCTGTTCCTTTGGGAAGTCGAAGTGTGAGAATTACA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        856 GAAAAAAGAGAAGAAAATGTAATTGCTGTTCCTTTGGGAAGTCGAAGTGTGAGAATTACA 915

Qy        781 GTGAAAGTACCTGCCTACCTCTTATTGAATCAAAAAACACTTCAAGGAAGCAAAGGAGAA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        916 GTGAAAGTACCTGCCTACCTCTTATTGAATCAAAAAACACTTCAAGGAAGCAAAGGAGAA 975

Qy        841 CACAGCTTTAACAGCCCCGGCCTCTTTGTCGTAGAAAACACAACAGTGGAATTTCAGAGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        976 CACAGCTTTAACAGCCCCGGCCTCTTTGTCGTAGAAAACACAACAGTGGAATTTCAGAGC 1035

Qy        901 GGCTCCGAGAGGCAAACTTTTAAGATTCCAGGACCTCTGATGGCTGATTCCATCTTCAAG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1036 GGCTCCGAGAGGCAAACTTTTAAGATTCCAGGACCTCTGATGGCTGATTCCATCTTCAAG 1095

Qy        961 ACCAGGTACACTGCAGCCAAAGACAGCGTGGTTCAGTTCTTCTTTTACCAGCCCATCAGT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1096 ACCAGGTACACTGCAGCCAAAGACAGCGTGGTTCAGTTCTTCTTTTACCAGCCCATCAGT 1155

Qy       1021 CATCAGTGGAGACAAACTGACTTCTTTCCCTGCACTGTGACGTGTGGAGGAGGTTATCAT 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1156 CATCAGTGGAGACAAACTGACTTCTTTCCCTGCACTGTGACGTGTGGAGGAGGTTATCAT 1215

Qy       1081 CTCAATTCTGCTGAATGTGTGGATATCCGCTTGAAGAGGGTAGTTCCTGACCATTATTGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1216 CTCAATTCTGCTGAATGTGTGGATATCCGCTTGAAGAGGGTAGTTCCTGACCATTATTGT 1275

Qy       1141 CACTACTACCCTGAAAATGTAAAACCAAAACCAAAACTGAAGGAATGCAGCATGGATCCC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1276 CACTACTACCCTGAAAATGTAAAACCAAAACCAAAACTGAAGGAATGCAGCATGGATCCC 1335

Qy       1201 TGCCCATCAAGTGATGGATTTAAAGAGATAATGCCCTATGACCACTTCCAACCTCTTCCT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1336 TGCCCATCAAGTGATGGATTTAAAGAGATAATGCCCTATGACCACTTCCAACCTCTTCCT 1395

Qy       1261 CGCTGGGAACATAATCCTTGGACTGCATGTTCCGTGTCCTGTGGAGGAGGGATTCAGAGA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1396 CGCTGGGAACATAATCCTTGGACTGCATGTTCCGTGTCCTGTGGAGGAGGGATTCAGAGA 1455

Qy       1321 CGGAGCTTTGTGTGTGTAGAGGAATCCATGCATGGAGAGATATTGCAGGTGGAAGAATGG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1456 CGGAGCTTTGTGTGTGTAGAGGAATCCATGCATGGAGAGATATTGCAGGTGGAAGAATGG 1515

Qy       1381 AAGTGCATGTACGCACCAAACCCAAGGTTATGCAAACTTGTAATCTGTTTGATTGCCCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1516 AAGTGCATGTACGCACCAAACCCAAGGTTATGCAAACTTGTAATCTGTTTGATTGCCCC 1575

Qy       1441 AAGTGGATTGCCATGGAGTGGTCTCAGTGCACAGTGACTTGTGGCCGAGGGTTACGGTAC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1576 AAGTGGATTGCCATGGAGTGGTCTCAGTGCACAGTGACTTGTGGCCGAGGGTTACGGTAC 1635

Qy       1501 CGGGTTGTTCTGTGTATTAACCACCGCGGAGAGCATGTTGGGGGCTGCAATCCACAACTG 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1636 CGGGTTGTTCTGTGTATTAACCACCGCGGAGAGCATGTTGGGGGCTGCAATCCACAACTG 1695

Qy       1561 AAGTTACACATCAAAGAAGAATGTGTCATTCCCATCCCGTGTTATAAACCAAAAGAAAAA 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1696 AAGTTACACATCAAAGAAGAATGTGTCATTCCCATCCCGTGTTATAAACCAAAAGAAAAA 1755

Qy       1621 AGTCCAGTGGAAGCAAAATTGCCTTGGCTGAAACAAGCACAAGAACTAGAAGAGACCAGA 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1756 AGTCCAGTGGAAGCAAAATTGCCTTGGCTGAAACAAGCACAAGAACTAGAAGAGACCAGA 1815

Qy       1681 ATAGCAACAGAAGAACCAACGTTCATTCCAGAACCCTGGTCAGCCTGCAGTACCACGTGT 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1816 ATAGCAACAGAAGAACCAACGTTCATTCCAGAACCCTGGTCAGCCTGCAGTACCACGTGT 1875

Qy       1741 GGGCCAGGTGTGCAGGTCCGTGAGGTGAAGTGCCGTGTGCTCCTCACATTCACGCAGACT 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1876 GGGCCAGGTGTGCAGGTCCGTGAGGTGAAGTGCCGTGTGCTCCTCACATTCACGCAGACT 1935

Qy       1801 GAGACTGAGCTGCCTGAGGAAGAGTGTGAAGGCCCCAAGCTGCCCACCGAACGGCCCTGC 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1936 GAGACTGAGCTGCCTGAGGAAGAGTGTGAAGGCCCCAAGCTGCCCACCGAACGGCCCTGC 1995

Qy       1861 CTCCTGGAAGCATGTGATGAGAGCCCGGCCTCCCGAGAGCTAGACATCCCTCTCCCTGAG 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1996 CTCCTGGAAGCATGTGATGAGAGCCCGGCCTCCCGAGAGCTAGACATCCCTCTCCCTGAG 2055

Qy       1921 GACAGTGAGACGACTTACGACTGGGATTACGCTGGGTTCACCCCTTGCACAGCAACATGC 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2056 GACAGTGAGACGACTTACGACTGGGATTACGCTGGGTTCACCCCTTGCACAGCAACATGC 2115

Qy       1981 TTGGGAGGCCATCAAGAAGCCACAGACAGCGTGGTTACATATCCAGACCAGCAGACAGTC 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2116 TTGGGAGGCCATCAAGAAGCCACAGACAGCGTGGTTACATATCCAGACCAGCAGACAGTC 2175

Qy       2041 AATGACAGCTTGTGTGATATGGTTCACCGTCCTCCAGCCATCAGCCAGCCCTGTAACACA 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2176 AATGACAGCTTGTGTGATATGGTTCACCGTCCTCCAGCCATCAGCCAGCCCTGTAACACA 2235

Qy       2101 GAGCCCTGTCCCCCTAGGTGGCATGTGGGCTCTTTGGGGGCCTGCTCAGCTACCTGTGGA 2160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2236 GAGCCCTGTCCCCCTAGGTGGCATGTGGGCTCTTTGGGGGCCTGCTCAGCTACCTGTGGA 2295

Qy       2161 GTTGGAATTCAGACCCGAGATGTGTACTGCCCTGCACCCAGGGGAGACCCTGCCCCTGCT 2220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2296 GTTGGAATTCAGACCCGAGATGTGTACTGCCCTGCACCCAGGGGAGACCCTGCCCCTGCT 2355

Qy       2221 GAGGAGTGCCGAGATGAAAAGCCCCATGCTTTACAAGCATGCAATCAGTTTGACTGCCCT 2280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2356 GAGGAGTGCCGAGATGAAAAGCCCCATGCTTTACAAGCATGCAATCAGTTTGACTGCCCT 2415

Qy       2281 CCTGGCTGGCACATTGAAGAATGGCAGCAGTGTTCCAGGACTTGTGGGGGGGGAACTCAG 2340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2416 CCTGGCTGGCACATTGAAGAATGGCAGCAGTGTTCCAGGACTTGTGGGGGGGGAACTCAG 2475

Qy       2341 AACAGAAGAGTCACCTGTCGGCAGCTGCTAACGGATGGCAGCTTTTTGAATCTCTCAGAT 2400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2476 AACAGAAGAGTCACCTGTCGGCAGCTGCTAACGGATGGCAGCTTTTTGAATCTCTCAGAT 2535

Qy       2401 GAATTGTGCCAAGGACCCAAGGCATCGTCTCACAAGTCCTGTGGCAGGACAGACTGTCCT 2460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2536 GAATTGTGCCAAGGACCCAAGGCATCGTCTCACAAGTCCTGTGGCAGGACAGACTGTCCT 2595

Qy       2461 CCACATTTACCTGTGGGAGACTGCTGGAAGTGTCCTGTCAGTTTTGGTTGGAATCCAG 2520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2596 CCACATTTAGCTGTGGGACACTGGTCGAAGTGTTCTGTCAGTTGTGGTGTTGGAATCCAG 2655

Qy    2521 AGAAGAAAGCAGGTGTGTCAAAGGCTGGCAGCCAAAGG-CGGCGCATCCCCCTCAGTGAG 2580
           |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db    2656 AGAAGAAAGCAGGTGTGTCAAAGGCTGGCAGCCAAAGGTCGGCGCATCCCCCTCAGTGAG 2715

Qy    2581 ATGATGTGCAGGGATCTACCAGGGTTCCCTCTTGTAAGATCTTGCCAGATGCCTGAGTGC 2640
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2716 ATGATGTGCAGGGATCTACCAGGGTTCCCTCTTGTAAGATCTTGCCAGATGCCTGAGTGC 2775

Qy    2641 AGTAAAATCAAATCAGAGATGAAGACAAAACTTGGTGAGCAGGGTCCGCAGATCCTCAGT 2700
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2776 AGTAAAATCAAATCAGAGATGAAGACAAAACTTGGTGAGCAGGGTCCGCAGATCCTCAGT 2835

Qy    2701 GTCCAGAGAGTCTACATTCAGACAAGGGAAGAGAAGCGTATTAACCTGACCATTGGTAGC 2760
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2836 GTCCAGAGAGTCTACATTCAGACAAGGGAAGAGAAGCGTATTAACCTGACCATTGGTAGC 2895

Qy    2761 AGAGCCTATTTGCTGCCCAACACATCCGTGATTATTAAGTGCCCCGTGCGACGATTCCAG 2820
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2896 AGAGCCTATTTGCTGCCCAACACATCCGTGATTATTAAGTGCCCCGTGCGACGATTCCAG 2955

Qy    2821 AAATCTCTGATCCAGTGGGAGAAGGATGGCCGTTGCCTGCAGAACTCCAAACGGCTTGGC 2880
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2956 AAATCTCTGATCCAGTGGGAGAAGGATGGCCGTTGCCTGCAGAACTCCAAACGGCTTGGC 3015

Qy    2881 ATCACCAAGTCAGGCTCACTAAAAATCCACGGTCTTGCTGCCCCCGACATCGGCGTGTAC 2940
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3016 ATCACCAAGTCAGGCTCACTAAAAATCCACGGTCTTGCTGCCCCCGACATCGGCGTGTAC 3075

Qy    2941 CGGTGCATTGCAGGCTCTGCACAGGAAACAGTTGTGCTCAAGCTCATTGGTACTGACAAC 3000
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3076 CGGTGCATTGCAGGCTCTGCACAGGAAACAGTTGTGCTCAAGCTCATTGGTACTGACAAC 3135

Qy    3001 CGGCTCATCGCACGCCCAGCCCTCAGGGAGCCTATGAGGGAATATCCTGGGATGGACCAC 3060
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3136 CGGCTCATCGCACGCCCAGCCCTCAGGGAGCCTATGAGGGAATATCCTGGGATGGACCAC 3195

Qy    3061 AGCGAAGCCAATAGTTTGGGAGTCACATGGCACAAAATGAGGCAAATGTGGAATAACAAA 3120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3196 AGCGAAGCCAATAGTTTGGGAGTCACATGGCACAAAATGAGGCAAATGTGGAATAACAAA 3255

Qy    3121 AATGACCTTTATCTGGATGATGACCACATTAGTAACCAGCCTTTCTTGAGAGCTCTGTTA 3180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3256 AATGACCTTTATCTGGATGATGACCACATTAGTAACCAGCCTTTCTTGAGAGCTCTGTTA 3315

Qy    3181 GGCCACTGCAGCAATTCTGCAGGAAGCACCAACTCCTGGGAGTTGAAGAATAAGCAGTTT 3240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3316 GGCCACTGCAGCAATTCTGCAGGAAGCACCAACTCCTGGGAGTTGAAGAATAAGCAGTTT 3375

Qy    3241 GAAGCAGCAGTTAAACAAGGAGCATATAGCATGGATACAGCCCAGTTTGATGAGCTGATA 3300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3376 GAAGCAGCAGTTAAACAAGGAGCATATAGCATGGATACAGCCCAGTTTGATGAGCTGATA 3435

Qy    3301 AGAAACATGAGTCAGCTCATGGAAACCGGAGAGGTCAGCGATGATCTTGCGTCCCAGCTG 3360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3436 AGAAACATGAGTCAGCTCATGGAAACCGGAGAGGTCAGCGATGATCTTGCGTCCCAGCTG 3495

Qy    3361 ATATATCAGCTGGTGGCCGAATTAGCCAAGGCACAGCCAACACACATGCAGTGGCGGGGC 3420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3496 ATATATCAGCTGGTGGCCGAATTAGCCAAGGCACAGCCAACACACATGCAGTGGCGGGGC 3555

Qy    3421 ATCCAGGAAGAGACACCTCCTGCTGCTCAGCTCAGAGGGGAAACAGGGAGTGTGTCCCAA 3480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3556 ATCCAGGAAGAGACACCTCCTGCTGCTCAGCTCAGAGGGGAAACAGGGAGTGTGTCCCAA 3615

Qy    3481 AGCTCGCATGCAAAAAACTCAGGCAAGCTGACATTCAAGCCGAAAGGACCTGTTCTCATG 3540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3616 AGCTCGCATGCAAAAAACTCAGGCAAGCTGACATTCAAGCCGAAAGGACCTGTTCTCATG 3675

Qy    3541 AGGCAAAGCCAACCTCCCTCAATTTCATTTAATAAAACAATAAATTCCAGGATTGGAAAT 3600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3676 AGGCAAAGCCAACCTCCCTCAATTTCATTTAATAAAACAATAAATTCCAGGATTGGAAAT 3735
```

```
Qy    3601 ACAGTATACATTACAAAAAGGACAGAGGTCATCAATATACTGTGTGACCTTATTACCCCC 3660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3736 ACAGTATACATTACAAAAAGGACAGAGGTCATCAATATACTGTGTGACCTTATTACCCCC 3795

Qy    3661 AGTGAGGCCACATATACATGGACCAAGGATGGAACCTTGTTACAGCCCTCAGTAAAAATA 3720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3796 AGTGAGGCCACATATACATGGACCAAGGATGGAACCTTGTTACAGCCCTCAGTAAAAATA 3855

Qy    3721 ATTTTGGATGGAACTGGGAAGATACAGATACAGAATCCTACAAGGAAAGAACAAGGCATA 3780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3856 ATTTTGGATGGAACTGGGAAGATACAGATACAGAATCCTACAAGGAAAGAACAAGGCATA 3915

Qy    3781 TATGAATGTTCTGTAGCTAATCATCTTGGTTCAGATGTGGAAAGTTCTTCTGTGCTGTAT 3840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3916 TATGAATGTTCTGTAGCTAATCATCTTGGTTCAGATGTGGAAAGTTCTTCTGTGCTGTAT 3975

Qy    3841 GCAGAGGCACCTGTCATCTTGTCTGTTGAAAGAAATATCACCAAACCAGAGCACAACCAT 3900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3976 GCAGAGGCACCTGTCATCTTGTCTGTTGAAAGAAATATCACCAAACCAGAGCACAACCAT 4035

Qy    3901 CTGTCTGTTGTGGTTGGAGGCATCGTGGAGGCAGCCCTTGGAGCAAACGTGACAATCCGA 3960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4036 CTGTCTGTTGTGGTTGGAGGCATCGTGGAGGCAGCCCTTGGAGCAAACGTGACAATCCGA 4095

Qy    3961 TGTCCTGTAAAAGGTGTCCCTCAGCCTAATATAACTTGGTTGAAGAGAGGAGGATCTCTG 4020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4096 TGTCCTGTAAAAGGTGTCCCTCAGCCTAATATAACTTGGTTGAAGAGAGGAGGATCTCTG 4155

Qy    4021 AGTGGCAATGTTTCCTTGCTTTTCAATGGATCCCTGTTGTTGCAGAATGTTTCCCTTGAA 4080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4156 AGTGGCAATGTTTCCTTGCTTTTCAATGGATCCCTGTTGTTGCAGAATGTTTCCCTTGAA 4215

Qy    4081 AATGAAGGAACCTACGTCTGCATAGCCACCAATGCTCTTGGAAAGGCAGTGGCAACATCT 4140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4216 AATGAAGGAACCTACGTCTGCATAGCCACCAATGCTCTTGGAAAGGCAGTGGCAACATCT 4275

Qy    4141 GTACTCCACTTGCTGGAACGAAGATGGCCAGAGAGTAGAATCGTATTTCTGCAAGGACAT 4200
           ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4276 GTATTCCACTTGCTGGAACGAAGATGGCCAGAGAGTAGAATCGTATTTCTGCAAGGACAT 4335

Qy    4201 AAAAAGTACATTCTCCAGGCAACCAACACTAGAACCAACAGCAATGACCCAACAGGAGAA 4260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4336 AAAAAGTACATTCTCCAGGCAACCAACACTAGAACCAACAGCAATGACCCAACAGGAGAA 4395

Qy    4261 CCCCCGCCTCAAGAGCCTTTTTGGGAGCCTGGTAACTGGTCACATTGTTCTGCCACCTGT 4320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4396 CCCCCGCCTCAAGAGCCTTTTTGGGAGCCTGGTAACTGGTCACATTGTTCTGCCACCTGT 4455

Qy    4321 GGTCATTTGGGAGCCCGCATTCAGAGACCCCAGTGTGTGATGGCCAATGGGCAGGAAGTG 4380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4456 GGTCATTTGGGAGCCCGCATTCAGAGACCCCAGTGTGTGATGGCCAATGGGCAGGAAGTG 4515

Qy    4381 AGTGAGGCCCTGTGTGATCACCTCCAGAAGCCACTGGCTGGGTTTGAGCCCTGTAACATC 4440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4516 AGTGAGGCCCTGTGTGATCACCTCCAGAAGCCACTGGCTGGGTTTGAGCCCTGTAACATC 4575

Qy    4441 CGGGACTGCCCAGCGAGGTGGTTCACAAGTGTGTGGTCACAGTGCTCTGTGTCTTGCGGT 4500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4576 CGGGACTGCCCAGCGAGGTGGTTCACAAGTGTGTGGTCACAGTGCTCTGTGTCTTGCGGT 4635

Qy    4501 GAAGGATACCACAGTCGGCAGGTGACGTGCAAGCGGACAAAAGCCAATGGAACTGTGCAG 4560
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4636 GAAGGATACCACAGTCGGCAGGTGACGTGCAAGCGGACAAAAGCCAATGGAACTGTGCAG 4695

Qy    4561 GTGGTGTCTCCAAGAGCATGTGCCCCTAAAGACCGGCCTCTGGGAAGAAAACCATGTTTT 4620
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4696 GTGGTGTCTCCAAGAGCATGTGCCCCTAAAGACCGGCCTCTGGGAAGAAAACCATGTTTT 4755

Qy    4621 GGTCATCCATGTGTTCAGTGGGAACCAGGGAACCGGTGTCCTGGACGTTGCATGGCCGT 4680
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4756 GGTCATCCATGTGTTCAGTGGGAACCAGGGAACCGGTGTCCTGGACGTTGCATGGCCGT 4815
```

```
Qy       4681 GCTGTGAGGATGCAGCAGCTGTACACAGCTTGTCAACACAACAGCTCTGACTCCAACTCT 4740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4816 GCTGTGAGGATGCAGCAGCTGTACACAGCTTGTCAACACAACAGCTCTGACTCCAACTCT 4875

Qy       4741 GATGACAGAAAGAGACCCACCTTAAGAAGGAACTGCACATCAGGGACCTTTGATGCTGCT 4800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4876 GATGACAGAAAGAGACCCACCTTAAGAAGGAACTGCACATCAGGGACCTTTGATGCTGCT 4935

Qy       4801 TGGCACACAGGCCCTGGGAAGCCCTGTACAGCAGCCTGTGGCAGGGCTTCCAGTCTCCTG 4860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4936 TGGCACACAGGCCCTGGGAAGCCCTGTACAGCAGCCTGTGGCAGGGCTTCCAGTCTCCTG 4995

Qy       4861 AAAGTCGACTGTATCCACACAAGGAGTTGCAAACCTGTGGCCAAGAGACACTCTGTACAT 4920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4996 AAAGTCGACTGTATCCACACAAGGAGTTGCAAACCTGTGGCCAAGAGACACTCTGTACAT 5055

Qy       4921 AAAAAGAAACCAATTTCCTGGCGGCACTGTCTTGGGCCCTCCTGTGATAGAGACTGCACA 4980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5056 AAAAAGAAACCAATTTCCTGGCGGCACTGTCTTGGGCCCTCCTGTGATAGAGACTGCACA 5115

Qy       4981 GACACAACTCACTACTGTATGTTTGTAAAACATCTAATTTGTGTTCTTAGACCTGCTACT 5040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5116 GACACAACTCACTACTGTATGTTTGTAAAACATCTAATTTGTGTTCTTAGACCTGCTACT 5175

Qy       5041 AAACAAAGGTGCTGCCAGTCATCTCAAGAGGGATAA 5076
              ||||||||||||||||||||||||||||||||||||
Db       5176 AAACAAAGGTGCTGCCAGTCATCTCAAGAGGGATAA 5211


RESULT 8
AX224823
LOCUS       AX224823                4854 bp    DNA     linear   PAT 10-SEP-2001
DEFINITION  Sequence 15 from Patent WO0161011.
ACCESSION   AX224823
VERSION     AX224823.1  GI:15554916
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 4854)
  AUTHORS   Donoho,G., Scoville,J., Turner,C.A., Friedrich,G., Zambrowicz,B.
            and Sands,A.T.
  TITLE     Novel human thrombospondin repeat proteins and polynucleotides
            encoding the same
  JOURNAL   Patent: WO 0161011-A 15 23-AUG-2001;
            Lexicon Genetics Incorporated (US)
FEATURES             Location/Qualifiers
     source          1..4854
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT     1331 a   1175 c   1276 g   1072 t
ORIGIN

  Query Match             93.9%;  Score 4766;  DB 6;  Length 4854;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 4766;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        311 CTGGAAGGAATTGTGAAGGGCAGAACATTCGGTACAAGACATGCAGCAATCATGACTGCC 370
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         89 CTGGAAGGAATTGTGAAGGGCAGAACATTCGGTACAAGACATGCAGCAATCATGACTGCC 148

Qy        371 CTCCAGATGCAGAAGATTTCAGACCCAGCAGTGCTCAGCCTACAATGATGTCCCAGTATC 430
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        149 CTCCAGATGCAGAAGATTTCAGACCCAGCAGTGCTCAGCCTACAATGATGTCCCAGTATC 208

Qy        431 AGGGGCATTACTATGAATGGCTTCCACGATATAATGATCCTGCTGCCCCCTGTGGACTCA 490
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        209 AGGGGCATTACTATGAATGGCTTCCACGATATAATGATCCTGCTGCCCCCTGTGGACTCA 268

Qy        491 AGTGTCATGCACAAGGACAAAATTGGTGGTGGAGCTGGCACCTAAGGTACTGGATGGGAA 550
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        269 AGTGTCATGCACAAGGACAAAATTGGTGGTGGAGCTGGCACCTAAGGTACTGGATGGGAA 328

Qy        551 CTGGTTGCAACACGGACTCCTTGGACATGTGTATCAGTGGCATCTGTCAGGCAGTGGGCT 610
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        329 CTGGTTGCAACACGGACTCCTTGGACATGTGTATCAGTGGCATCTGTCAGGCAGTGGGCT 388

Qy        611 GCGATCGGCAACTGGGAAGCAATGCCAAGGAGGACAACTGTGGAGTCTGTGCCGGCGATG 670
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        389 GCGATCGGCAACTGGGAAGCAATGCCAAGGAGGACAACTGTGGAGTCTGTGCCGGCGATG 448

Qy        671 GCTCCACCTGCAGGCTTGTACGGGGACAATCAAAGTCACACGTTTCTCCTGAAAAAAGAG 730
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        449 GCTCCACCTGCAGGCTTGTACGGGGACAATCAAAGTCACACGTTTCTCCTGAAAAAAGAG 508

Qy        731 AAGAAAATGTAATTGCTGTTCCTTTGGGAAGTCGAAGTGTGAGAATTACACTGAAAGGAC 790
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        509 AAGAAAATGTAATTGCTGTTCCTTTGGGAAGTCGAAGTGTGAGAATTACACTGAAAGGAC 568

Qy        791 CTGCCCACCTCTTTATTGAATCAAAAACACTTCAAGGAAGCAAAGGAGAACACAGCTTTA 850
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        569 CTGCCCACCTCTTTATTGAATCAAAAACACTTCAAGGAAGCAAAGGAGAACACAGCTTTA 628

Qy        851 ACAGCCCCGGCGTCTTTGTCGTAGAAAACACAACAGTGGAATTTCAGAGGGGCTCCGAGA 910
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        629 ACAGCCCCGGCGTCTTTGTCGTAGAAAACACAACAGTGGAATTTCAGAGGGGCTCCGAGA 688

Qy        911 GGCAAACTTTTAAGATTCCAGGACCTCTGATGGCTGATTTCATCTTCAAGACCAGGTACA 970
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        689 GGCAAACTTTTAAGATTCCAGGACCTCTGATGGCTGATTTCATCTTCAAGACCAGGTACA 748

Qy        971 CTGCAGCCAAAGACAGCGTGGTTCAGTTCTTCTTTTACCAGCCCATCAGTCATCAGTGGA 1030
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        749 CTGCAGCCAAAGACAGCGTGGTTCAGTTCTTCTTTTACCAGCCCATCAGTCATCAGTGGA 808

Qy       1031 GACAAACTGACTTCTTTCCCTGCACTGTGACGTGTGGAGGAGGTTATCAGCTCAATTCTG 1090
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        809 GACAAACTGACTTCTTTCCCTGCACTGTGACGTGTGGAGGAGGTTATCAGCTCAATTCTG 868

Qy       1091 CTGAATGTGTGGATATCCGCTTGAAGAGGGTAGTTCCTGACCATTATTGTCACTACTACC 1150
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        869 CTGAATGTGTGGATATCCGCTTGAAGAGGGTAGTTCCTGACCATTATTGTCACTACTACC 928

Qy       1151 CTGAAAATGTAAAACCAAAACCAAAACTGAAGGAATGCAGCATGGATCCCTGCCCATCAA 1210
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        929 CTGAAAATGTAAAACCAAAACCAAAACTGAAGGAATGCAGCATGGATCCCTGCCCATCAA 988

Qy       1211 GTGATGGATTTAAAGAGATAATGCCCTATGACCACTTCCAACCTCTTCCTCGCTGGGAAC 1270
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        989 GTGATGGATTTAAAGAGATAATGCCCTATGACCACTTCCAACCTCTTCCTCGCTGGGAAC 1048

Qy       1271 ATAATCCTTGGACTGCATGTTCCGTGTCCTGTGGAGGAGGGATTCAGAGACGGAGCTTTG 1330
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1049 ATAATCCTTGGACTGCATGTTCCGTGTCCTGTGGAGGAGGGATTCAGAGACGGAGCTTTG 1108

Qy       1331 TGTGTGTAGAGGAATCCATGCATGGAGAGATATTGCAGGTGGAAGAATGGAAGTGCATGT 1390
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1109 TGTGTGTAGAGGAATCCATGCATGGAGAGATATTGCAGGTGGAAGAATGGAAGTGCATGT 1168

Qy       1391 ACGCACCCAAACCCAAGGTTATGCAAACTTGTAATCTGTTTGATTGCCCCAAGTGGATTG 1450
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1169 ACGCACCCAAACCCAAGGTTATGCAAACTTGTAATCTGTTTGATTGCCCCAAGTGGATTG 1228

Qy       1451 CCATGGAGTGGTCTCAGTGCACAGTGACTTGTGGCCGAGGGTTACGGTACCGGGTTGTTC 1510
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1229 CCATGGAGTGGTCTCAGTGCACAGTGACTTGTGGCCGAGGGTTACGGTACCGGGTTGTTC 1288

Qy       1511 TGTGTATTAACCACCGCGGAGAGCATGTTGGGGCTGCAATCCTACAACTGAAGTTACACA 1570
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1289 TGTGTATTAACCACCGCGGAGAGCATGTTGGGGCTGCAATCCTACAACTGAAGTTACACA 1348

Qy       1571 CTAAAGAAGAATGTGTCATTCCCATCCGTGTTATAAACCAAAAGAAAAAAGTCCCAGTGG 1630
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1349 CTAAAGAAGAATGTGTCATTCCCATCCGTGTTATAAACCAAAAGAAAAAAGTCCCAGTGG 1408
```

```
Qy       1631 AAGCAAAATTGCCTTGGCTGAAACAAGCACAAGAACTAGAAGAGACCAGAATAGCAACAG 1690
              .|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1409 AAGCAAAATTGCCTTGGCTGAAACAAGCACAAGAACTAGAAGAGACCAGAATAGCAACAG 1468

Qy       1691 AAGAACCAACGTTCATTCCAGAACCCTGGTCAGCCTGCAGTACCACGTGTGGGCCAGGTG 1750
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1469 AAGAACCAACGTTCATTCCAGAACCCTGGTCAGCCTGCAGTACCACGTGTGGGCCAGGTG 1528

Qy       1751 TGCAGGTCCGCGAGGTGAAGTGCCGTGTGCTCCTCACATTCACGCAGACTGAGACTGAGC 1810
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1529 TGCAGGTCCGCGAGGTGAAGTGCCGTGTGCTCCTCACATTCACGCAGACTGAGACTGAGC 1588

Qy       1811 TGCCCGAGGAAGAGTGTGAAGGCCCCAAGCTGCCCACCGAACGGCCCTGCCTCCTGGAAG 1870
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1589 TGCCCGAGGAAGAGTGTGAAGGCCCCAAGCTGCCCACCGAACGGCCCTGCCTCCTGGAAG 1648

Qy       1871 CATGTGATGAGAGCCCGGCCTCCCGAGAGCTAGACATCCCTCTCCCTGAGGACAGTGAGA 1930
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1649 CATGTGATGAGAGCCCGGCCTCCCGAGAGCTAGACATCCCTCTCCCTGAGGACAGTGAGA 1708

Qy       1931 CGACTTACGACTGGGAGTACGCTGGGTTCACCCCTTGCACAGCAACATGCTTGGGAGGCC 1990
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1709 CGACTTACGACTGGGAGTACGCTGGGTTCACCCCTTGCACAGCAACATGCTTGGGAGGCC 1768

Qy       1991 ATCAAGAAGCCATAGCAGTGTGCTTACATATCCAGACCCAGCAGACAGTCAATGACAGCT 2050
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1769 ATCAAGAAGCCATAGCAGTGTGCTTACATATCCAGACCCAGCAGACAGTCAATGACAGCT 1828

Qy       2051 TGTGTGATATGGTCCACCGTCCTCCAGCCATGAGCCAGGCCTGTAACACAGAGCCCTGTC 2110
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1829 TGTGTGATATGGTCCACCGTCCTCCAGCCATGAGCCAGGCCTGTAACACAGAGCCCTGTC 1888

Qy       2111 CCCCCAGGTGGCATGTGGGCTCTTGGGGGCCCTGCTCAGCTACCTGTGGAGTTGGAATTC 2170
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1889 CCCCCAGGTGGCATGTGGGCTCTTGGGGGCCCTGCTCAGCTACCTGTGGAGTTGGAATTC 1948

Qy       2171 AGACCCGAGATGTGTACTGCCTGCACCCAGGGGAGACCCCTGCCCCTCCTGAGGAGTGCC 2230
              ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db       1949 AGACCCGAGATGTGTACTGCCTGCACCCAGGGGAGACCCCTGCCCCTCCTGAGGAGTGCC 2008

Qy       2231 GAGATGAAAAGCCCCATGCTTTACAAGCATGCAATCAGTTTGACTGCCCTCCTGGCTGGC 2290
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2009 GAGATGAAAAGCCCCATGCTTTACAAGCATGCAATCAGTTTGACTGCCCTCCTGGCTGGC 2068

Qy       2291 ACATTGAAGAATGGCAGCAGTGTTCCAGGACTTGTGGCGGGGGAACTCAGAACAGAAGAG 2350
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2069 ACATTGAAGAATGGCAGCAGTGTTCCAGGACTTGTGGCGGGGGAACTCAGAACAGAAGAG 2128

Qy       2351 TCACCTGTCGGCAGCTGCTAACGGATGGCAGCTTTTTGAATCTCTCAGATGAATTGTGCC 2410
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2129 TCACCTGTCGGCAGCTGCTAACGGATGGCAGCTTTTTGAATCTCTCAGATGAATTGTGCC 2188

Qy       2411 AAGGACCCAAGGCATCGTCTCACAAGTCCTGTGCCAGGACAGACTGTCCTCCACATTTAG 2470
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2189 AAGGACCCAAGGCATCGTCTCACAAGTCCTGTGCCAGGACAGACTGTCCTCCACATTTAG 2248

Qy       2471 CTGTGGGAGACTGGTCGAAGTGTTCTGTCAGTTGTGGTGTTGGAATCCAGAGAAGAAAGC 2530
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2249 CTGTGGGAGACTGGTCGAAGTGTTCTGTCAGTTGTGGTGTTGGAATCCAGAGAAGAAAGC 2308

Qy       2531 AGGTGTGTCAAAGGCTGGCAGCCAAAGGTCGGCGCATCCCCCTCAGTGAGATGATGTGCA 2590
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2309 AGGTGTGTCAAAGGCTGGCAGCCAAAGGTCGGCGCATCCCCCTCAGTGAGATGATGTGCA 2368

Qy       2591 GGGATCTACCAGGGTTCCCTCTTGTAAGATCTTGCCAGATGCCTGAGTGCAGTAAAATCA 2650
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2369 GGGATCTACCAGGGTTCCCTCTTGTAAGATCTTGCCAGATGCCTGAGTGCAGTAAAATCA 2428

Qy       2651 AATCAGAGATGAAGACAAAACTTGGTGAGCAGGGTCCGCAGATCCTCAGTGTCCAGAGAG 2710
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2429 AATCAGAGATGAAGACAAAACTTGGTGAGCAGGGTCCGCAGATCCTCAGTGTCCAGAGAG 2488

Qy       2711 TCTACATTCAGACAAGGGAAGAGAAGCGTATTAACCTGACCATTGGTAGCAGAGCCTATT 2770
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2489 TCTACATTCAGACAAGGGAAGAGAAGCGTATTAACCTGACCATTGGTAGCAGAGCCTATT 2548

Qy       2771 TGCTGCCCAACACATCCGTGATTATTAAGTGCCCCGTGCGACGATTCCAGAAATCTCTGA 2830
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2549 TGCTGCCCAACACATCCGTGATTATTAAGTGCCCCGTGCGACGATTCCAGAAATCTCTGA 2608

Qy       2831 TCCAGTGGGAGAAGGATGGCCGTTGCCTGCAGAACTCCAAACGGCTTGGCATCACCAAGT 2890
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2609 TCCAGTGGGAGAAGGATGGCCGTTGCCTGCAGAACTCCAAACGGCTTGGCATCACCAAGT 2668

Qy       2891 CAGGCTCACTAAAAATCCACGGTCTTGCTGCCCCCGACATCGGCGTGTACCGGTGCATTG 2950
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2669 CAGGCTCACTAAAAATCCACGGTCTTGCTGCCCCCGACATCGGCGTGTACCGGTGCATTG 2728

Qy       2951 CAGGCTCTGCACAGGAAACAGTTGTGCTCAAGCTCATTGGTACTGACAACCGGCTCATCG 3010
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2729 CAGGCTCTGCACAGGAAACAGTTGTGCTCAAGCTCATTGGTACTGACAACCGGCTCATCG 2788

Qy       3011 CACGCCCAGCCCTCAGGGAGCCTATGAGGGAATATCCTGGGATGGACCACAGCGAAGCCA 3070
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2789 CACGCCCAGCCCTCAGGGAGCCTATGAGGGAATATCCTGGGATGGACCACAGCGAAGCCA 2848

Qy       3071 ATAGTTTGGGAGTCACATGGCACAAAATGAGGCAAATGTGGAATAACAAAAATGACCTTT 3130
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2849 ATAGTTTGGGAGTCACATGGCACAAAATGAGGCAAATGTGGAATAACAAAAATGACCTTT 2908

Qy       3131 ATCTGGATGATGACCACATTAGTAACCAGCCTTTCTTGAGAGCTCTGTTAGGCCACTGCA 3190
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2909 ATCTGGATGATGACCACATTAGTAACCAGCCTTTCTTGAGAGCTCTGTTAGGCCACTGCA 2968

Qy       3191 GCAATTCTGCAGGAAGCACCAACTCCTGGGAGTTGAAGAATAAGCAGTTTGAAGCAGCAG 3250
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2969 GCAATTCTGCAGGAAGCACCAACTCCTGGGAGTTGAAGAATAAGCAGTTTGAAGCAGCAG 3028

Qy       3251 TTAAACAAGGAGCATATAGCATGGATACAGCCCAGTTTGATGAGCTGATAAGAAACATGA 3310
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3029 TTAAACAAGGAGCATATAGCATGGATACAGCCCAGTTTGATGAGCTGATAAGAAACATGA 3088

Qy       3311 GTCAGCTCATGGAAACCGGAGAGGTCAGCCGATGATCTTGCGTCCCAGCTGATATATCAGC 3370
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3089 GTCAGCTCATGGAAACCGGAGAGGTCAGCCGATGATCTTGCGTCCCAGCTGATATATCAGC 3148

Qy       3371 TGGTGGCCGAATTAGCCAAGGCACAGCCAACACACATGCAGTGGCGGGGCATCCAGGAAG 3430
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3149 TGGTGGCCGAATTAGCCAAGGCACAGCCAACACACATGCAGTGGCGGGGCATCCAGGAAG 3208

Qy       3431 AGACACCTCCTGCTGCTCAGCTCAGAGGGGAAACAGGGAGTGTGTCCCAAAGCTCGCATG 3490
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3209 AGACACCTCCTGCTGCTCAGCTCAGAGGGGAAACAGGGAGTGTGTCCCAAAGCTCGCATG 3268

Qy       3491 CAAAAAACTCAGGCAAGCTGACATTCAAGCCGAAAGGACCTGTTCTCATGAGGCAAAGCC 3550
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3269 CAAAAAACTCAGGCAAGCTGACATTCAAGCCGAAAGGACCTGTTCTCATGAGGCAAAGCC 3328

Qy       3551 AACCTCCCTCAATTTCATTTAATAAAACAATAAATTCCAGGATTGGAAATACAGTATACA 3610
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3329 AACCTCCCTCAATTTCATTTAATAAAACAATAAATTCCAGGATTGGAAATACAGTATACA 3388

Qy       3611 TTACAAAAAGGACAGAGGTCATCAATATACTGTGTGACCTTATTACCCCCAGTGAGGCCA 3670
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3389 TTACAAAAAGGACAGAGGTCATCAATATACTGTGTGACCTTATTACCCCCAGTGAGGCCA 3448

Qy       3671 CATATACATGGACCAAGGATGGAACCTTGTTACAGCCCTCAGTAAAAATAATTTTGGATG 3730
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3449 CATATACATGGACCAAGGATGGAACCTTGTTACAGCCCTCAGTAAAAATAATTTTGGATG 3508

Qy       3731 GAACTGGGAAGATACAGATACAGAATCCTACAAGGAAAGAACAAGGCATATATGAATGTT 3790
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3509 GAACTGGGAAGATACAGATACAGAATCCTACAAGGAAAGAACAAGGCATATATGAATGTT 3568

Qy       3791 CTGTAGCTAATCATCTTGGTTCAGATGTGGAAAGTTCTTCTGTGCTGTATGCAGAGGCAC 3850
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3569 CTTTAGCTAATCATTTTTTTAGATGTGGAAAGTTCTTCTGTCCTCTGTATGAGAGCCCAC 3628

Qy       3851 CTGTCATCTTGTCTGTTGAAAGAAATATCACCAAACCAGAGCACAACCATCTGTCTGTTG 3910
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3629 CTGTCATCTTGTCTGTTGAAAGAAATATCACCAAACCAGAGCACAACCATCTGTCTGTTG 3688

Qy       3911 TGGTTGGAGGCATCGTGGAGGCAGCCCTTGGAGCAAACGTGACAATCCGATGTCCTGTAA 3970
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3689 TGGTTGGAGGCATCGTGGAGGCAGCCCTTGGAGCAAACGTGACAATCCGATGTCCTGTAA 3748

Qy       3971 AAGGTGTCCCTCAGCCTAATATAATTTGGTTGAAGAGAGGAGGATCTCTGAGTGGCAATG 4030
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3749 AAGGTGTCCCTCAGCCTAATATAATTTGGTTGAAGAGAGGAGGATCTCTGAGTGGCAATG 3808

Qy       4031 TTTCCTTGCTTTTCAATGGATCCCTGTTGTTGCAGAATGTTTCCCTTGAAAATGAAGGAA 4090
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3809 TTTCCTTGCTTTTCAATGGATCCCTGTTGTTGCAGAATGTTTCCCTTGAAAATGAAGGAA 3868

Qy       4091 CCTACGTCTGCATAGCCACCAATGCTCTTGGAAAGGCAGTGGCAACATCTGTACTCCACT 4150
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3869 CCTACGTCTGCATAGCCACCAATGCTCTTGGAAAGGCAGTGGCAACATCTGTACTCCACT 3928

Qy       4151 TGCTGGAACGAAGATGGCCAGAGAGTAGAATCGTATTTCTGCAAGGACATAAAAAGTACA 4210
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3929 TGCTGGAACGAAGATGGCCAGAGAGTAGAATCGTATTTCTGCAAGGACATAAAAAGTACA 3988

Qy       4211 TTCTCCAGGCAACCAACACTAGAACCAACAGCAATGACCCAACAGGAGAACCCCCGCCTC 4270
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3989 TTCTCCAGGCAACCAACACTAGAACCAACAGCAATGACCCAACAGGAGAACCCCCGCCTC 4048

Qy       4271 AAGAGCCTTTTTGGGAGCCTGGTAACTGGTCACATTGTTCTGCCACCTGTGGTCATTTGG 4330
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4049 AAGAGCCTTTTTGGGAGCCTGGTAACTGGTCACATTGTTCTGCCACCTGTGGTCATTTGG 4108

Qy       4331 GAGCCCGCATTCAGAGACCCCAGTGTGTGATGGCCAATGGGCAGGAAGTGAGTGAGGCCC 4390
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4109 GAGCCCGCATTCAGAGACCCCAGTGTGTGATGGCCAATGGGCAGGAAGTGAGTGAGGCCC 4168

Qy       4391 TGTGTGATCACCTCCAGAAGCCACTGGCTGGGTTTGAGCCCTGTAACATCCGGGACTGCC 4450
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4169 TGTGTGATCACCTCCAGAAGCCACTGGCTGGGTTTGAGCCCTGTAACATCCGGGACTGCC 4228

Qy       4451 CAGCGAGGTGGTTCACAAGTGTGTGGTCACAGTGCTCTGTGTCTTGCGGTGAAGGATACC 4510
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4229 CAGCGAGGTGGTTCACAAGTGTGTGGTCACAGTGCTCTGTGTCTTGCGGTGAAGGATACC 4288

Qy       4511 ACAGTCGGCAGGTGACGTGCAAGCGGACAAAAGCCAATGGAACTGTGCAGGTGGTGTCTC 4570
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4289 ACAGTCGGCAGGTGACGTGCAAGCGGACAAAAGCCAATGGAACTGTGCAGGTGGTGTCTC 4348

Qy       4571 CAAGAGCATGTGCCCCTAAAGACCGGCCTCTGGGAAGAAAACCATGTTTTGGTCATCCAT 4630
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4349 CAAGAGCATGTGCCCCTAAAGACCGGCCTCTGGGAAGAAAACCATGTTTTGGTCATCCAT 4408

Qy       4631 GTGTTCAGTGGGAACCAGGGAACCGGTGTCCTGGACGTTGCATGGGCCGTGCTGTGAGGA 4690
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4409 GTGTTCAGTGGGAACCAGGGAACCGGTGTCCTGGACGTTGCATGGGCCGTGCTGTGAGGA 4468

Qy       4691 TGCAGCAGCGTCACACAGCTTGTCAACACAACAGCTCTGACTCCAACTGTGATGACAGAA 4750
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4469 TGCAGCAGCGTCACACAGCTTGTCAACACAACAGCTCTGACTCCAACTGTGATGACAGAA 4528

Qy       4751 AGAGACCCACCTTAAGAAGGAACTGCACATCAGGGGCCTGTGATGTGTGTTGGCACACAG 4810
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4529 AGAGACCCACCTTAAGAAGGAACTGCACATCAGGGGCCTGTGATGTGTGTTGGCACACAG 4588

Qy       4811 GCCCTTGGAAGCCCTGTACAGCAGCCTGTGGCAGGGGTTTCCAGTCTCGGAAAGTCGACT 4870
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4589 GCCCTTGGAAGCCCTGTACAGCAGCCTGTGGCAGGGGTTTCCAGTCTCGGAAAGTCGACT 4648

Qy       4871 GTATCCACACAAGGAGTTGCAAACCTGTGGCCAAGAGACACTGTGTACAGAAAAAGAAAC 4930
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4649 GTATCCACACAAGGAGTTGCAAACCTGTGGCCAAGAGACACTGTGTACAGAAAAAGAAAC 4708

Qy       4931 CAATTTCCTGGCGGCACTGTCTTGGGCCTCTCTGTGATAGAGACTGCACAGACACAACTC 4990
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4709 CAATTTCCTGGCGGCACTGTCTTGGGCCTCTCTGTGATAGAGACTGCACAGACACAACTC 4768

Qy       4991 ACTACTGTATGTTTGTAAAACATCTTAATTTGTGTTCTCTAGACCGCTACAAACAAAGGT 5050
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4769 ACTACTGTATGTTTGTAAAACATCTTAATTTGTGTTCTCTAGACCGCTACAAACAAAGGT 4828

Qy       5051 GCTGCCAGTCATGTCAAGAGGGATAA 5076
              ||||||||||||||||||||||||||
Db       4829 GCTGCCAGTCATGTCAAGAGGGATAA 4854


RESULT 4
AX235410
LOCUS       AX235410                 6303 bp    DNA     linear   PAT 11-SEP-2001
DEFINITION  Sequence 57 from Patent WO0162928.
ACCESSION   AX235410
VERSION     AX235410.1  GI:15593924
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 6303)
  AUTHORS   Vernet,C.A., Fernandes,E., Shimkets,R.A., Macdougall,J. and
            Spaderna,S.K.
  TITLE     Polypeptides and nucleic acids encoding same
  JOURNAL   Patent: WO 0162928-A 57 30-AUG-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1..6303
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT     1792 a   1297 c   1528 g   1686 t
ORIGIN

  Query Match             76.1%;  Score 3862.2;  DB 6;  Length 6303;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 3864;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1210 AGTGATGGATTTAAAGAGATAATGCCCTATGACCACTTCCAACCTCTTCCTCGCTGGGAA 1269
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        404 AGTGATGGATTTAAAGAGATAATGCCCTATGACCACTTCCAACCTCTTCCTCGCTGGGAA 463

Qy       1270 CATAATCCTTGGACTGCATGTTTCCGTGTCCTGTGGAGGAGGATTCAGAGACGGAGCTTT 1329
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        464 CATAATCCTTGGACTGCATGTTTCCGTGTCCTGTGGAGGAGGATTCAGAGACGGAGCTTT 523

Qy       1330 GTGTGTGTAGAGGAATCCATGCATGGAGAGATATTGCAGGTGGAAGAATGGAAGTGCATG 1389
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        524 GTGTGTGTAGAGGAATCCATGCATGGAGAGATATTGCAGGTGGAAGAATGGAAGTGCATG 583

Qy       1390 TACGCACCCAAACCCAAGGTTATGCAAACTTGTAATCTGTTTGATTGCCCCAAGTGGATT 1449
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        584 TACGCACCCAAACCCAAGGTTATGCAAACTTGTAATCTGTTTGATTGCCCCAAGTGGATT 643

Qy       1450 GCCATGGAGTGGTCTCAGTGCACAGTGACTTGTGGCCGAGGGTTACGGTACCGGGTTGTT 1509
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        644 GCCATGGAGTGGTCTCAGTGCACAGTGACTTGTGGCCGAGGGTTACGGTACCGGGTTGTT 703

Qy       1510 CTGGTATTAACCACCGCGGGAGAGCATGTTGGGGGCTGCAATCCACAACTGAACTTACAC 1569
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        704 CTGGTATTAACCACCGCGGGAGAGCATGTTGGGGGCTGCAATCCACAACTGAACTTACAC 763

Qy       1570 ATCAAAGAAGAATGTGTCATTCCCATCCCGTGTTATAAACCAAAAGAAAAAAGTCCAGTG 1629
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        764 ATCAAAGAAGAATGTGTCATTCCCATCCCGTGTTATAAACCAAAAGAAAAAAGTCCAGTG 823

Qy       1630 GAAGCAAAATTGCCTTGGCTGAAACAAGCACAAGAACTAGAAGAGACCAGAATAGCAACA 1689
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        824 GAAGCAAAATTGCCTTGGCTGAAACAAGCACAAGAACTAGAAGAGACCAGAATAGCAACA 883
```

```
Qy       1690 GAAGAACCAACGTTCATTCCAGAACCCTGGTCAGCCTGCAGTACCACGTGTGGGCCAGGT 1749
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db        884 GAAGAACCAACGTTCATTCCAGAACCCTGGTCAGCCTGCAGTACCACGTGTGGGCCGGGT 943

Qy       1750 GTGCAGGTCCGCGAGGTGAAGTGCCGTGTGCTCCTCACATTCACGCAGACTGAGACTGAG 1809
              ||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        944 GTGCAGGTCCGTGAGGTGAAGTGCCGTGTGCTCCTCACATTCACGCAGACTGAGACTGAG 1003

Qy       1810 CTGCCCGAGGAAGAGTGTGAAGGCCCCAAGCTGCCCACCGAACGGCCCTGCCTCCTGGAA 1869
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1004 CTGCCCGAGGAAGAGTGTGAAGGCCCCAAGCTGCCCACCGAACGGCCCTGCCTCCTGGAA 1063

Qy       1870 GCATGTGATGAGAGCCCGGCCTCCCGAGAGCTAGACATCCCTCTCCCTGAGGACAGTGAG 1929
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1064 GCATGTGATGAGAGCCCGGCCTCCCGAGAGCTAGACATCCCTCTCCCTGAGGACAGTGAG 1123

Qy       1930 ACGACTTACGACTGGAGTACGCTGGGTTCACCCCTTGCACAGCAACATGCTTGGGAGGC 1989
              ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db       1124 ACGACTTACGACTGGAGTACGCTGGGTTCACCCCTTGCACAGCAACATGCGTGGGAGGC 1183

Qy       1990 CATCAAGAAGCCATAGCAGTGTGCTTACATATCCAGACCCAGCAGACAGTCAATGACAGC 2049
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1184 CATCAAGAAGCCATAGCAGTGTGCTTACATATCCAGACCCAGCAGACAGTCAATGACAGC 1243

Qy       2050 TTGTGTGATATGGTCCACCGTCCTCCAGCCATGAGCCAGGCCTGTAACACAGAGCCCTGT 2109
              ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db       1244 TTGTGTGATATGGTCCACCGTCCTCCAGCCATGAGCCAGGCCTGTAACACAGAGCCCTGT 1303

Qy       2110 CCCCCAGGTGGCATGTGGGCTCTTGGGGGCCCTGCTCAGCTACCTGTGGAGTTGGAATT 2169
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1304 CCCCCAGGTGGCATGTGGGCTCTTGGGGGCCCTGCTCAGCTACCTGTGGAGTTGGAATT 1363

Qy       2170 CAGACCCGAGATGTGTACTGCCTGCACCCAGGGGAGACCCCTGCCCCTCCTGAGGACTGC 2229
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1364 CAGACCCGAGATGTGTACTGCCTGCACCCAGGGGAGACCCCTGCCCCTCCTGAGGACTGC 1423

Qy       2230 CGAGATGAAAAGCCCCATGCTTTACAAGCATGCAATCAGTTTGACTGCCCTCCTGGCTGG 2289
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1424 CGAGATGAAAAGCCCCATGCTTTACAAGCATGCAATCAGTTTGACTGCCCTCCTGGCTGG 1483

Qy       2290 CACATTGAAGAATGGCAGCAGTGTTCCAGGACTTGTGGCGGGGGAACTCAGAACAGAAGA 2349
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1484 CACATTGAAGAATGGCAGCAGTGTTCCAGGACTTGTGGCGGGGGAACTCAGAACAGAAGA 1543

Qy       2350 GTCACCTGTCGGCAGCTGCTAACGGATGGCAGCTTTTTGAATCTCTCAGATGAATTGTGC 2409
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1544 GTCACCTGTCGGCAGCTGCTAACGGATGGCAGCTTTTTGAATCTCTCAGATGAATTGTGC 1603

Qy       2410 CAAGGACCCAAGGCATCGTCTCACAAGTCCTGTGCCAGGACAGACTGTCCTCCACATTTA 2469
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1604 CAAGGACCCAAGGCATCGTCTCACAAGTCCTGTGCCAGGACAGACTGTCCTCCACATTTA 1663

Qy       2470 GCTGTGGGAGACTGGTCGAAGTGTTCTGTCAGTTGTGGTGTTGGAATCCAGAGAAGAAAG 2529
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1664 GCTGTGGGAGACTGGTCGAAGTGTTCTGTCAGTTGTGGTGTTGGAATCCAGAGAAGAAAG 1723

Qy       2530 CAGGTGTGTCAAAGGCTGGCAGCCAAAGGTCGGCGCATCCCCCTCAGTGAGATGATGTGC 2589
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1724 CAGGTGTGTCAAAGGCTGGCAGCCAAAGGTCGGCGCATCCCCCTCAGTGAGATGATGTGC 1783

Qy       2590 AGGGATCTACCAGGGTTCCCTCTTGTAAGATCTTGCCAGATGCCTGAGTGCAGTAAAATC 2649
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1784 AGGGATCTACCAGGGTTCCCTCTTGTAAGATCTTGCCAGATGCCTGAGTGCAGTAAAATC 1843

Qy       2650 AAATCAGAGATGAAGACAAAACTTGGTGAGCAGGGTCCGCAGATCCTCAGTGTCCAGAGA 2709
              ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db       1844 AAATCAGAGATGAAGACAAAACTTGGTGAGCAGGGTCCGCAGATCCTCAGTGTCCAGAGA 1903

Qy       2710 GTCTACATTCAGACAAGGGAAGAGAAGCGTATTAACCTGACCATTGGTAGCAGAGCCTAT 2769
              |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db       1904 GTCTACATTCAGACAAGGGAAGAGAAGCGTATTAACCTGACCATTGGTAGCAGAGCCTAT 1963

Qy       2770 TTGCTGCCCAACACATCCGTGATTATTAAGTGCCCCGTGCGACGATTCCAGAAATCTCTG 2829
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1964 TTGCTGCCCAACACATCCGTGATTATTAAGTGCCCCGTGCGACGATTCCAGAAATCTCTG 2023

Qy       2830 ATCCAGTGGGAGAAGGATGGCCGTTGCCTGCAGAACTCCAAACGGCTTGGCATCACCAAG 2889
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2024 ATCCAGTGGGAGAAGGATGGCCGTTGCCTGCAGAACTCCAAACGGCTTGGCATCACCAAG 2083

Qy       2890 TCAGGCTCACTAAAAATCCACGGTCTTGCTGCCCCCGACATCGGCGTGTACCGGTGCATT 2949
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2084 TCAGGCTCACTAAAAATCCACGGTCTTGCTGCCCCCGACATCGGCGTGTACCGGTGCATT 2143

Qy       2950 GCAGGCTCTGCACAGGAAACAGTTGTGCTCAAGCTCATTGGTACTGACAACCGGCTCATC 3009
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2144 GCAGGCTCTGCACAGGAAACAGTTGTGCTCAAGCTCATTGGTACTGACAACCGGCTCATC 2203

Qy       3010 GCACGCCCAGCCCTCAGGGAGCCTATGAGGGAATATCCTGGGATGGACCACAGCGAAGCC 3069
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2204 GCACGCCCAGCCCTCAGGGAGCCTATGAGGGAATATCCTGGGATGGACCACAGCGAAGCC 2263

Qy       3070 AATAGTTTGGGAGTCACATGGCACAAAATGAGGCAAATGTGGAATAACAAAAATGACCTT 3129
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2264 AATAGTTTGGGAGTCACATGGCACAAAATGAGGCAAATGTGGAATAACAAAAATGACCTT 2323

Qy       3130 TATCTGGATGATGACCACATTAGTAACCAGCCTTTCTTGAGAGCTCTGTTAGGCCACTGC 3189
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2324 TATCTGGATGATGACCACATTAGTAACCAGCCTTTCTTGAGAGCTCTGTTAGGCCACTGC 2383

Qy       3190 AGCAATTCTGCAGGAAGCACCAACTCCTGGGAGTTGAAGAATAAGCAGTTTGAAGCAGCA 3249
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2384 AGCAATTCTGCAGGAAGCACCAACTCCTGGGAGTTGAAGAATAAGCAGTTTGAAGCAGCA 2443

Qy       3250 GTTAAACAAGGAGCATATAGCATGGATACAGCCCAGTTTGATGAGCTGATAAGAAACATG 3309
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2444 GTTAAACAAGGAGCATATAGCATGGATACAGCCCAGTTTGATGAGCTGATAAGAAACATG 2503

Qy       3310 AGTCAGCTCATGGAAACCGGAGAGGTCAGCGATGATCTTGCGTCCCAGCTGATATATCAG 3369
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2504 AGTCAGCTCATGGAAACCGGAGAGGTCAGCGATGATCTTGCGTCCCAGCTGATATATCAG 2563

Qy       3370 CTGGTGGCCGAATTAGCCAAGGCACAGCCAACACACATGCAGTGGCGGGGCATCCAGGAA 3429
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2564 CTGGTGGCCGAATTAGCCAAGGCACAGCCAACACACATGCAGTGGCGGGGCATCCAGGAA 2623

Qy       3430 GAGACACCTCCTGCTGCTCAGCTCAGAGGGGAAACAGGGAGTGTGTCCCAAAGCTCGCAT 3489
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2624 GAGACACCTCCTGCTGCTCAGCTCAGAGGGGAAACAGGGAGTGTGTCCCAAAGCTCGCAT 2683

Qy       3490 GCAAAAAACTCAGGCAAGCTGACATTCAAGCCGAAAGGACCTGTTCTCATGAGGCAAAGC 3549
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2684 GCAAAAAACTCAGGCAAGCTGACATTCAAGCCGAAAGGACCTGTTCTCATGAGGCAAAGC 2743

Qy       3550 CAACCTCCCTCAATTTCATTTAATAAAACAATAAATTCCAGGATTGGAAATACAGTATAC 3609
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2744 CAACCTCCCTCAATTTCATTTAATAAAACAATAAATTCCAGGATTGGAAATACAGTATAC 2803

Qy       3610 ATTACAAAAAGGACAGAGGTCATCAATATACTGTGTGACCTTATTACCCCCAGTGAGGCC 3669
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2804 ATTACAAAAAGGACAGAGGTCATCAATATACTGTGTGACCTTATTACCCCCAGTGAGGCC 2863

Qy       3670 ACATATACATGGACCAAGGATGGAACCTTGTTACAGCCCTCAGTAAAAATAATTTTGGAT 3729
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2864 ACATATACATGGACCAAGGATGGAACCTTGTTACAGCCCTCAGTAAAAATAATTTTGGAT 2923

Qy       3730 GGAACTGGGAAGATACAGATACAGAATCCTACAAGGAAAGAACAAGGCATATATGAATGT 3789
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2924 GGAACTGGGAAGATACAGATACAGAATCCTACAAGGAAAGAACAAGGCATATATGAATGT 2983

Qy       3790 TCTGTAGCTAATCATCTTGGTTCAGATGTGGAAAGTTCTTCTGTGCTGTATGCAGAGGCA 3849
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2984 TCTGTAGCTAATCATCTTGGTTCAGATGTGGAAAGTTCTTCTGTGCTGTATGCAGAGGCA 3043

Qy       3850 CCTGTCATCTTGTCTGTTGAAAGAAATATCACCAAACCAGAGCACAACCATCTGTCTGTT 3909
```

```
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3044 CCTGTCATCTTGTCTGTTGAAAGAAATATCACCAAACCAGAGCACAACCATCTGTCTGTT 3103

Qy       3910 GTGGTTGGAGGCATCGTGGAGGCAGCCCTTGGAGCAAACGTGACAATCCGATGTCCTGTA 3969
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3104 GTGGTTGGAGGCATCGTGGAGGCAGCCCTTGGAGCAAACGTGACAATCCGATGTCCTGTA 3163

Qy       3970 AAAGGTGTCCCTCAGCCTAATATAACTTGGTTGAAGAGAGGAGGATCTCTGAGTGGCAAT 4029
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3164 AAAGGTGTCCCTCAGCCTAATATAACTTGGTTGAAGAGAGGAGGATCTCTGAGTGGCAAT 3223

Qy       4030 GTTTCCTTGCTTTTCAATGGATCCCTGTTGTTGCAGAATGTTTCCCTTGAAAATGAAGGA 4089
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3224 GTTTCCTTGCTTTTCAATGGATCCCTGTTGTTGCAGAATGTTTCCCTTGAAAATGAAGGA 3283

Qy       4090 ACCTACGTCTGCATAGCCACCAATGCTCTTGGAAAGGCAGTGGCAACATCTGTACTCCAC 4149
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3284 ACCTACGTCTGCATAGCCACCAATGCTCTTGGAAAGGCAGTGGCAACATCTGTACTCCAC 3343

Qy       4150 TTGCTGGAACGAAGATGGCCAGAGAGTAGAATCGTATTTCTGCAAGGACATAAAAAGTAC 4209
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3344 TTGCTGGAACGAAGATGGCCAGAGAGTAGAATCGTATTTCTGCAAGGACATAAAAAGTAC 3403

Qy       4210 ATTCTTCAGGCAACTAACACTAGAACCAACAGCAATGACCCAACAGGAGAACCCCCGCCT 4269
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3404 ATTCTTCAGGCAACTAACACTAGAACCAACAGCAATGACCCAACAGGAGAACCCCCGCCT 3463

Qy       4270 CAAGAGCCTTTTTGGGAGCCTGGTAACTGGTCACATTGTTCTGCCACCTGTGGTCATTTG 4329
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3464 CAAGAGCCTTTTTGGGAGCCTGGTAACTGGTCACATTGTTCTGCCACCTGTGGTCATTTG 3523

Qy       4330 GGAGCCCGGATTCAGAGACCCAGTGTGTGATGGCCAATGGGCAGGAAGTGAGTGAGGCCC 4389
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3524 GGAGCCCGGATTCAGAGACCCAGTGTGTGATGGCCAATGGGCAGGAAGTGAGTGAGGCCC 3583

Qy       4390 CTGTGTGATCACCTCCAGAAGCCACTGGCTGGGTTGAGCCCTGTAACATCCGGGACTGGC 4449
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3584 CTGTGTGATCACCTCCAGAAGCCACTGGCTGGGTTGAGCCCTGTAACATCCGGGACTGGC 3643

Qy       4450 CCAGCGAGGTGGTTCACAAGTGTGTGGTCACAGTGCTCTGTGTCTTGCGGTGAAGGATAC 4509
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3644 CCAGCGAGGTGGTTCACAAGTGTGTGGTCACAGTGCTCTGTGTCTTGCGGTGAAGGATAC 3703

Qy       4510 CACAGTCGGCAGGTGACGTGCAAGCGGACAAAAGCCAATGGAACTGTGCAGGTGGTGTCT 4569
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3704 CACAGTCGGCAGGTGACGTGCAAGCGGACAAAAGCCAATGGAACTGTGCAGGTGGTGTCT 3763

Qy       4570 CCAAGAGCATGTGCCCCTAAAGACCGGCCTCTGGGAAGAAAACCATGTTTTGGTCATCCA 4629
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3764 CCAAGAGCATGTGCCCCTAAAGACCGGCCTCTGGGAAGAAAACCATGTTTTGGTCATCCA 3823

Qy       4630 TGTGTTCAGTGGGAACCAGGGAACCGGTGTCCTGGACGTTGCATGGGCCGTGCTGTGAGG 4689
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3824 TGTGTTCAGTGGGAACCAGGGAACCGGTGTCCTGGACGTTGCATGGGCCGTGCTGTGAGG 3883

Qy       4690 ATGCAGCAGCGTCACACAGCTTGTCAACACAACAGCTCTGACTCCAACTGTGATGACAGA 4749
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3884 ATGCAGCAGCGTCACACAGCTTGTCAACACAACAGCTCTGACTCCAACTGTGATGACAGA 3943

Qy       4750 AAGAGACCCACCTTAAGAAGGAACTGCACATCAGGGGCCTGTGATGTGTGTTGGCACACA 4809
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3944 AAGAGACCCACCTTAAGAAGGAACTGCACATCAGGGGCCTGTGATGTGTGTTGGCACACA 4003

Qy       4810 GGCCCTTGGAAGCCCTGTACAGCAGCCTGTGGCAGGGGTTTCCAGTCTCGGAAAGTCGAC 4869
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4004 GGCCCTTGGAAGCCCTGTACAGCAGCCTGTGGCAGGGGTTTCCAGTCTCGGAAAGTCGAC 4063

Qy       4870 TGTATCCACACAAGGAGTTGCAAACCTGTGGCCAAGAGACACTGTGTACAGAAAAAGAAA 4929
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4064 TGTATCCACACAAGGAGTTGCAAACCTGTGGCCAAGAGACACTGTGTACAGAAAAAGAAA 4123

Qy       4930 CCAATTCCTGGGGGCACTGTCTTGGGCCTCCTGTGATAGACACTGCACACACACACAACT 4989
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4124 CCAATTCCTGGGGGCACTGTCTTGGGCCTCCTGTGATAGACACTGCACACACACACAACT 4183

Qy       4990 CACTACTGTATGTTTGTAAAACATCTTAATTTGTGTTCTCTAGACCGCTACAAACAAAGG 5049
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4184 CACTACTGTATGTTTGTAAAACATCTTAATTTGTGTTCTCTAGACCGCTACAAACAAAGG 4243

Qy       5050 TGCTGCCAGTCATGTCAAGAGGGATAA 5076
              |||||||||||||||||||||||||||
Db       4244 TGCTGCCAGTCATGTCAAGAGGGATAA 4270


RESULT 5
AX235354
LOCUS       AX235354                6294 bp    DNA     linear   PAT 11-SEP-2001
DEFINITION  Sequence 1 from Patent WO0162928.
ACCESSION   AX235354
VERSION     AX235354.1  GI:15593887
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 6294)
  AUTHORS   Vernet,C.A., Fernandes,E., Shimkets,R.A., Macdougall,J. and
            Spaderna,S.K.
  TITLE     Polypeptides and nucleic acids encoding same
  JOURNAL   Patent: WO 0162928-A 1 30-AUG-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1..6294
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             416..4261
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC69811.1"
                     /db_xref="GI:15593888"
                     /translation="MPYDHFQPLPRWEHNPWTACSVSCGGGIQRRSFVCVEESMHGEI
                     LQVEEWKCMYAPKPKVMQTCNLFDCPKWIAMEWSQCTVTCGRGLRYRVVLCINHRGEH
                     VGGCNPQLKLHIKEECVIPIPCYKPKEKSPVEAKLPWLKQAQELEETRIATEEPTFIP
                     EPWSACSTTCGPGVQVREVKCRVLLTFTQTETELPEEECEGPKLPTERPCLLEACDES
                     PASRELDIPLPEDSETTYDWEYAGFTPCTATCVGGHQEATAVCLHIQTQQTVNDSLCD
                     MVHRPPAMSQACNTEPCPPRWHVGSWGPCSATCGVGIQTRDVYCLHPGETPAPPEECR
                     DEKPHALQACNQFDCPPGWHIEEWQQCSRTCGGGTQNRRVTCRQLLTDGSFLNLSDEL
                     CQGPKASSHKSCARTDCPPHLAVGDWSKCSVSCGVGIQRRKQVCQRLAAKGRRIPLSE
                     MMCRDLPGLPLVRSCQMPECSKIKSEMKTKLGEQGPQILSVQRVYIQTREEKRINLTI
                     GSRAYLLPNTSVIIKCPVRRFQKSLIQWEKDGRCLQNSKRLGITKSGSLKIHGLAAPD
                     IGVYRCIAGSAQETVVLKLIGTDNRLIARPALREPMREYPGMDHSEANSLGVTWHKMR
                     QMWNNKNDLYLDDDHISNQPFLRALLGHCSNSAGSTNSWELKNKQFEAAVKQGAYSMD
                     TAQFDELIRNMSQLMETGEVSDDLASQLIYQLVAELAKAQPTHMQWRGIQEETPPAAQ
                     LRGETGSVSQSSHAKNSGKLTFKPKGPVLMRQSQPPSISFNKTINSRIGNTVYITKRT
                     EVINILCDLITPSEATYTWTKDGTLLQPSVKIILDGTGKIQIQNPTRKEQGIYECSVA
                     NHLGSDVESSSVLYAEAPVILSVERNITKPEHNHLSVVVGGIVEAALGANVTIRCPVK
                     GVPQPNITWLKRGGSLSGNVSLLFNGSLLLQNVSLENEGTYVCIATNALGKAVATSVL
                     HLLERRWPESRIVFLQGHKKYILQATNTRTNSNDPTGEPPPQEPFWEPGNWSHCSATC
                     GHLGARIQRPQCVMANGQEVSEALCDHLQKPLAGFEPCNIRDCPARWFTSVWSQCSVS
                     CGEGYHSRQVTCKRTKANGTVQVVSPRACAPKDRPLGRKPCFGHPCVQWEPGNRCPGR
                     CMGRAVRMQQRHTACQHNSSDSNCDDRKRPTLRRNCTSGACDVCWHTGPWRPCTAACG
                     RGFQSRKVDCIHTRSCKPVAKRHCVQKKKPISWRHCLGPSCDRDCTDTTHYCMFVKHL
                     NLCSLDRYKQRCCQSCQEG"
BASE COUNT     1789 a   1297 c   1528 g   1680 t
ORIGIN

  Query Match             76.0%;  Score 3859;  DB 6;  Length 6294;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 3862;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1210 AGTGATGGATTTAAAGAGATAATGCCCTATGACCACTTCCAACCTCTTCCTCGCTGGGAA 1269
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        395 AGTGATGGATTTAAAGAGATAATGCCCTATGACCACTTCCAACCTCTTCCTCGCTGGGAA 454

Qy       1270 CATAATCCTTGGACTGCATGTTCTGTGTCCTGTGGAGGAGGGATTCAGAGACGGAGCTTT 1329
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        455 CATAATCCTTGGACTGCATGTTCCGTGTCCTGTGGAGGAGGGATTCAGAGACGGAGCTTT 514

Qy       1330 GTGTGTGTAGAGGAATCCATGCATGGAGAGATATTGCAGGTGGAAGAATGGAAGTGCATG 1389
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        515 GTGTGTGTAGAGGAATCCATGCATGGAGAGATATTGCAGGTGGAAGAATGGAAGTGCATG 574

Qy       1390 TACGCACCCAAACCCAAGGTTATGCAAACTTGTAATCTGTTTGATTGCCCCAAGTGGATT 1449
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        575 TACGCACCCAAACCCAAGGTTATGCAAACTTGTAATCTGTTTGATTGCCCCAAGTGGATT 634

Qy       1450 GCCATGGAGTGGTCTCAGTGCACAGTGACTTGTGGCCGAGGGTTACGGTACCGGGTTGTT 1509
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        635 GCCATGGAGTGGTCTCAGTGCACAGTGACTTGTGGCCCAGGGTTACGGTACCGGGTTGTT 694

Qy       1510 CTGTGTATTAACCACCGCGGAGAGCATGTTGGGGGCTGCAATCCACAACTGAAGTTACAC 1569
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        695 CTGTGTATTAACCACCGCGGAGAGCATGTTGGGGGCTGCAATCCACAACTGAAGTTACAC 754

Qy       1570 ATCAAAGAAGAATGTGTCATTCCCATCCCGTGTTATAAACCAAAAGAAAAAAGTCCAGTG 1629
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        755 ATCAAAGAAGAATGTGTCATTCCCATCCCGTGTTATAAACCAAAAGAAAAAAGTCCAGTG 814

Qy       1630 GAAGCAAAATTGCCTTGGCTGAAACAAGCACAAGAACTAGAAGAGACCAGAATAGCAACA 1689
              ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        815 GAAGCAAAATTGCCTTGGCTGAAACAACCACAAGAACTAGAAGAGACCAGAATAGCAACA 874

Qy       1690 GAAGAACCAACGTTCATTCCAGAACCCTGGTCAGCCTGCAGTACCACGTGTGGGCCAGGT 1749
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db        875 GAAGAACCAACGTTCATTCCAGAACCCTGGTCAGCCTGCAGTACCACGTGTGGGCCGGGT 934

Qy       1750 GTGCAGGTCCGCGAGGTGAAGTGCCGTGTGCTCCTCACATTCACGCAGACTGAGACTGAG 1809
              ||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        935 GTGCAGGTCCGTGAGGTGAAGTGCCGTGTGCTCCTCACATTCACGCAGACTGAGACTGAG 994

Qy       1810 CTGCCCGAGGAAGAGTGTGAAGGCCCCAAGCTGCCCACCGAACGGCCCTGCCTCCTGGAA 1869
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        995 CTGCCCGAGGAAGAGTGTGAAGGCCCCAAGCTGCCCACCGAACGGCCCTGCCTCCTGGAA 1054

Qy       1870 GCATGTGATGAGAGCCCGGCCTCCCGAGAGCTAGACATCCCTCTCCCTGAGGACAGTGAG 1929
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1055 GCATGTGATGAGAGCCCGGCCTCCCGAGAGCTAGACATCCCTCTCCCTGAGGACAGTGAG 1114

Qy       1930 ACGACTTACGACTGGGAGTACGCTGGGTTCACCCCTTGCACAGCAACATGCTTGGGAGGC 1989
              ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db       1115 ACGACTTACGACTGGGAGTACGCTGGGTTCACCCCTTGCACAGCAACATGCGTGGGAGGC 1174

Qy       1990 CATCAAGAAGCCATAGCAGTGTGCTTACATATCCAGACCCAGCAGACAGTCAATGACAGC 2049
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1175 CATCAAGAAGCCATAGCAGTGTGCTTACATATCCAGACCCAGCAGACAGTCAATGACAGC 1234

Qy       2050 TTGTGTGATATGGTCCACCGTCCTCCAGCCATGAGCCAGGCCTGTAACACAGAGCCCTGT 2109
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1235 TTGTGTGATATGGTCCACCGTCCTCCAGCCATGAGCCAGGCCTGTAACACAGAGCCCTGT 1294

Qy       2110 CCCCCAGGTGGCATGTGGGCTCTTGGGGGCCCTGCTCAGCTACCTGTGGAGTTGGAATT 2169
              |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db       1295 CCCCCAGGTGGCATGTGGGCTCTTGGGGGCCCTGCTCAGCTACCTGTGGAGTTGGAATT 1354

Qy       2170 CAGACCCGAGATGTGTACTGCCTGCACCCAGGGAGACCCCTGCCCCTCCTGAGGAGTGC 2229
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1355 CAGACCCGAGATGTGTACTGCCTGCACCCAGGGAGACCCCTGCCCCTCCTGAGGAGTGC 1414

Qy       2230 CGAGATGAAAAGCCCCATGCTTTACAAGCATGCAATCAGTTTGACTGCCCTCCTGGCTGG 2289
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1415 CGAGATGAAAAGCCCCATGCTTTACAAGCATGCAATCAGTTTGACTGCCCTCCTGGCTGG 1474

Qy       2290 CACATTGAAGAATGGCAGCAGTGTTCCAGGACTTGTGGCGGGGGAACTCAGAACAGAAGA 2349
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1475 CACATTGAAGAATGGCAGCAGTGTTCCAGGACTTGTGGCGGGGGAACTCAGAACAGAAGA 1534

Qy       2350 GTCACCTGTCGGCAGCTGCTAACGGATGGCAGCTTTTTGAATCTCTCAGATGAATTGTGC 2409
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1535 GTCACCTGTCGGCAGCTGCTAACGGATGGCAGCTTTTTGAATCTCTCAGATGAATTGTGC 1594

Qy       2410 CAAGGACCCAAGGCATCGTCTCACAAGTCCTGTGCCAGGACAGACTGTCCTCCACATTTA 2469
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1595 CAAGGACCCAAGGCATCGTCTCACAAGTCCTGTGCCAGGACAGACTGTCCTCCACATTTA 1654

Qy       2470 GCTGTGGGAGACTGGTCGAAGTGTTCTGTCAGTTGTGGTGTTGGAATCCAGAGAAGAAAG 2529
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1655 GCTGTGGGAGACTGGTCGAAGTGTTCTGTCAGTTGTGGTGTTGGAATCCAGAGAAGAAAG 1714

Qy       2530 CAGGTGTGTCAAAGGCTGGCAGCCAAAGGTCGGCGCATCCCCCTCAGTGAGATGATGTGC 2589
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1715 CAGGTGTGTCAAAGGCTGGCAGCCAAAGGTCGGCGCATCCCCCTCAGTGAGATGATGTGC 1774

Qy       2590 AGGGATCTACCAGGGTTCCCTCTTGTAAGATCTTGCCAGATGCCTGAGTGCAGTAAAATC 2649
              ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db       1775 AGGGATCTACCAGGGCTCCCTCTTGTAAGATCTTGCCAGATGCCTGAGTGCAGTAAAATC 1834

Qy       2650 AAATCAGAGATGAAGACAAAACTTGGTGAGCAGGGTCCGCAGATCCTCAGTGTCCAGAGA 2709
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1835 AAATCAGAGATGAAGACAAAACTTGGTGAGCAGGGTCCGCAGATCCTCAGTGTCCAGAGA 1894

Qy       2710 GTCTACATTCAGACAAGGGAAGAGAAGCGTATTAACCTGACCATTGGTAGCAGAGCCTAT 2769
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1895 GTCTACATTCAGACAAGGGAAGAGAAGCGTATTAACCTGACCATTGGTAGCAGAGCCTAT 1954

Qy       2770 TTGCTGCCCAACACATCCGTGATTATTAAGTGCCCCGTGCGACGATTCCAGAAATCTCTG 2829
              ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db       1955 TTGCTGCCCAACACATCCGTGATTATTAAGTGCCCAGTGCGACGATTCCAGAAATCTCTG 2014

Qy       2830 ATCCAGTGGGAGAAGGATGGCCGTTGCCTGCAGAACTCCAAACGGCTTGGCATCACCAAG 2889
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2015 ATCCAGTGGGAGAAGGATGGCCGTTGCCTGCAGAACTCCAAACGGCTTGGCATCACCAAG 2074

Qy       2890 TCAGGCTCACTAAAAATCCACGGTCTTGCTGCCCCCGACATCGGCGTGTACCGGTGCATT 2949
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2075 TCAGGCTCACTAAAAATCCACGGTCTTGCTGCCCCCGACATCGGCGTGTACCGGTGCATT 2134

Qy       2950 GCAGGCTCTGCACAGGAAACAGTTGTGCTCAAGCTCATTGGTACTGACAACCGGCTCATC 3009
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2135 GCAGGCTCTGCACAGGAAACAGTTGTGCTCAAGCTCATTGGTACTGACAACCGGCTCATC 2194

Qy       3010 GCACGCCCAGCCCTCAGGGAGCCTATGAGGGAATATCCTGGGATGGACCACAGCGAAGCC 3069
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2195 GCACGCCCAGCCCTCAGGGAGCCTATGAGGGAATATCCTGGGATGGACCACAGCGAAGCC 2254

Qy       3070 AATAGTTTGGGAGTCACATGGCACAAAATGAGGCAAATGTGGAATAACAAAAATGACCTT 3129
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2255 AATAGTTTGGGAGTCACATGGCACAAAATGAGGCAAATGTGGAATAACAAAAATGACCTT 2314

Qy       3130 TATCTGGATGATGACCACATTAGTAACCAGCCTTTCTTGAGAGCTCTGTTAGGCCACTGC 3189
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2315 TATCTGGATGATGACCACATTAGTAACCAGCCTTTCTTGAGAGCTCTGTTAGGCCACTGC 2374

Qy       3190 AGCAATTCTGCAGGAAGCACCAACTCCTGGGAGTTGAAGAATAAGCAGTTTGAAGCAGCA 3249
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2375 AGCAATTCTGCAGGAAGCACCAACTCCTGGGAGTTGAAGAATAAGCAGTTTGAAGCAGCA 2434

Qy       3250 GTTAAACAAGGAGCATATAGCATGGATACAGCCCAGTTTGATGAGCTGATAAGAAACATG 3309
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2435 GTTAAACAAGGAGCATATAGCATGGATACAGCCCAGTTTGATGAGCTGATAAGAAACATG 2494

Qy       3310 AGTCAGCTCATGGAAACCGGAGAGGTCAGCGATGATCTTGCGTCCCAGCTGATATATCAG 3369
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2495 AGTCAGCTCATGGAAACCGGAGAGGTCAGCGATGATCTTGCGTCCCAGCTGATATATCAG 2554

Qy       3370 CTGGTGGCCCAATTAGCCAAGGCACAGCCAACACACATGCAGTGGCGGGGCATCCAGGAA 3429
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2555 CTGGTGGCCGAATTAGCCAAGGCACAGCCAACACACATGCAGTGGCGGGGCATCCAGGAA 2614

Qy       3430 GAGACACCTCCTGCTGCTCAGCTCAGAGGGGAAACAGGGAGTGTGTCCCAAAGCTCGCAT 3489
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2615 GAGACACCTCCTGCTGCTCAGCTCAGAGGGGAAACAGGGAGTGTGTCCCAAAGCTCGCAT 2674
```

```
Qy       3490 GCAAAAAACTCAGGCAAGCTGACATTCAAGCCGAAAGGACCTGTTCTCATGAGGCAAAGC 3549
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2675 GCAAAAAACTCAGGCAAGCTGACATTCAAGCCGAAAGGACCTGTTCTCATGAGGCAAAGC 2734

Qy       3550 CAACCTCCCTCAATTTCATTTAATAAAACAATAAATTCCAGGATTGGAAATACAGTATAC 3609
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2735 CAACCTCCCTCAATTTCATTTAATAAAACAATAAATTCCAGGATTGGAAATACAGTATAC 2794

Qy       3610 ATTACAAAAAGGACAGAGGTTCATCAATATACTGTGTGACCTTATTACCCCCAGTGAGGC 3669
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2795 ATTACAAAAAGGACAGAGGTTCATCAATATACTGTGTGACCTTATTACCCCCAGTGAGGC 2854

Qy       3670 ACATATACATGGACCAAGGATGGAACCTTGTTACAGCCCTCAGTAAAAATAATTTTGGAT 3729
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2855 ACATATACATGGACCAAGGATGGAACCTTGTTACAGCCCTCAGTAAAAATAATTTTGGAT 2914

Qy       3730 GGAACTGGGAAGATACAGATACAGAATCCTACAAGGAAAGAACAAGGCATATATGAATGT 3789
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2915 GGAACTGGGAAGATACAGATACAGAATCCTACAAGGAAAGAACAAGGCATATATGAATGT 2974

Qy       3790 TCTGTAGCTAATCATCTTGGTTCAGATGTGGAAAGTTCTTCTGTGCTGTATGCAGAGGCA 3849
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2975 TCTGTAGCTAATCATCTTGGTTCAGATGTGGAAAGTTCTTCTGTGCTGTATGCAGAGGCA 3034

Qy       3850 CCTGTCATCTTGTCTGTTGAAAGAAATATCACCAAACCAGAGCACAACCATCTGTCTGTT 3909
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3035 CCTGTCATCTTGTCTGTTGAAAGAAATATCACCAAACCAGAGCACAACCATCTGTCTGTT 3094

Qy       3910 GTGGTTGGAGGCATCGTGGAGGCAGCCCTTGGAGCAAACGTGACAATCCGATGTCCTGTA 3969
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3095 GTGGTTGGAGGCATCGTGGAGGCAGCCCTTGGAGCAAACGTGACAATCCGATGTCCTGTA 3154

Qy       3970 AAAGGTGTCCCTCAGCCTAATATAACTTGGTTGAAGAGAGGAGGATCTCTGAGTGGCAAT 4029
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3155 AAAGGTGTCCCTCAGCCTAATATAACTTGGTTGAAGAGAGGAGGATCTCTGAGTGGCAAT 3214

Qy       4030 GTTTCCTTGCTTTTCAATGGATCCCTGTTGTTGCAGAATGTTTCCCTTGAAAATGAAGGA 4089
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3215 GTTTCCTTGCTTTTCAATGGATCCCTGTTGTTGCAGAATGTTTCCCTTGAAAATGAAGGA 3274

Qy       4090 ACCTACGTCTGCATAGCCACCAATGCTCTTGGAAAGGCAGTGGCAACATCTGTACTCCAC 4149
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3275 ACCTACGTCTGCATAGCCACCAATGCTCTTGGAAAGGCAGTGGCAACATCTGTACTCCAC 3334

Qy       4150 TTGCTGGAACGAAGATGGCCAGAGAGTAGAATCGTATTTCTGCAAGGACATAAAAAGTAC 4209
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3335 TTGCTGGAACGAAGATGGCCAGAGAGTAGAATCGTATTTCTGCAAGGACATAAAAAGTAC 3394

Qy       4210 ATTCTCCAGGCAACCAACACTAGAACCAACAGCAATGACCCAACAGGAGAACCCCCGCCT 4269
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3395 ATTCTCCAGGCAACCAACACTAGAACCAACAGCAATGACCCAACAGGAGAACCCCCGCCT 3454

Qy       4270 CAAGAGCCTTTTTGGGAGCCTGGTAACTGGTCACATTGTTCTGCCACCTGTGGTCATTTG 4329
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3455 CAAGAGCCTTTTTGGGAGCCTGGTAACTGGTCACATTGTTCTGCCACCTGTGGTCATTTG 3514

Qy       4330 GGAGCCCGCATTCAGAGACCCCAGTGTGTGATGGCCAATGGGCAGGAAGTGAGTGAGGCC 4389
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3515 GGAGCCCGCATTCAGAGACCCCAGTGTGTGATGGCCAATGGGCAGGAAGTGAGTGAGGCC 3574

Qy       4390 CTGTGTGATCACCTCCAGAAGCCACTGGCTGGGTTTGAGCCCTGTAACATCCGGGACTGC 4449
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3575 CTGTGTGATCACCTCCAGAAGCCACTGGCTGGGTTTGAGCCCTGTAACATCCGGGACTGC 3634

Qy       4450 CCAGCGAGGTGGTTCACAAGTGTGTGGTCACAGTGCTCTGTGTCTTGGCGGTGAAGGATAC 4509
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3635 CCAGCGAGGTGGTTCACAAGTGTGTGGTCACAGTGCTCTGTGTCTTGGCGGTGAAGGATAC 3694

Qy       4510 CACAGTCGGCAGGTGACGTGCAAGCGGACAAAAGCCAATGGAACTGTGCAGGTGGTGTCT 4569
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3695 CACAGTCGGCAGGTGACGTGCAAGCGGACAAAAGCCAATGGAACTGTGCAGGTGGTGTCT 3754

Qy       4570 CCAAGAGCATGTGCCCCTAAAGACCGGCCTCTGGGAAGAAAAC ATGTTTGGTCATCCA 4629
              |||||||||||||||||||||||||||||||||||||||||||  |||||||||||||||
Db       3755 CCAAGAGCATGTGCCCCTAAAGACCGGCCTCTGGGAAGAAAAC ATGTTTGGTCATCCA 3814

Qy       4630 TGTGTTCAGTGGGAACCAGGGAACCGGGTGTCCTGGACGTTGCATGGGCCGTGCTGTGAGG 4689
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3815 TGTGTTCAGTGGGAACCAGGGAACCGGGTGTCCTGGACGTTGCATGGGCCGTGCTGTGAGG 3874

Qy       4690 ATGCAGCAGCGTCACACAGTTTGTTAACACAACAGCTCTGACTCCAACTGTGATGACAGA 4749
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3875 ATGCAGCAGCGTCACACAGTTTGTTAACACAACAGCTCTGACTCCAACTGTGATGACAGA 3934

Qy       4750 AAGAGACCCACCTTAAGAAGGAACTGCACATCAGGGGCTGTGATGTGTGTTGGCACACA 4809
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3935 AAGAGACCCACCTTAAGAAGGAACTGCACATCAGGGGCTGTGATGTGTGTTGGCACACA 3994

Qy       4810 GGCCCTTGGAAGCCCTGTACAGCAGCCTGTGGCAGGGGTTTCCAGTCTCGGAAAGTCGAC 4869
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3995 GGCCCTTGGAAGCCCTGTACAGCAGCCTGTGGCAGGGGTTTCCAGTCTCGGAAAGTCGAC 4054

Qy       4870 TGTATCCACACAAGGAGTTGCAAACCTGTGGCCAAGAGACACTGGTCTACAGAAAAAGAAA 4929
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4055 TGTATCCACACAAGGAGTTGCAAACCTGTGGCCAAGAGACACTGGTCTACAGAAAAAGAAA 4114

Qy       4930 CCAATTTCCTGGCGGCACTGTCTTGGGCCCTCCTGTGATAGAGACTGCACAGACACAACT 4989
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4115 CCAATTTCCTGGCGGCACTGTCTTGGGCCCTCCTGTGATAGAGACTGCACAGACACAACT 4174

Qy       4990 CACTACTGTATGTTTGTAAAACATCTTAATTTGTGTTCTCTAGACCGCTACAAACAAAGG 5049
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4175 CACTACTGTATGTTTGTAAAACATCTTAATTTGTGTTCTCTAGACCGCTACAAACAAAGG 4234

Qy       5050 TGCTGCCAGTCATGTCAAGAGGGATAA 5076
              |||||||||||||||||||||||||||
Db       4235 TGCTGCCAGTCATGTCAAGAGGGATAA 4261


RESULT 6
AX224825
LOCUS       AX224825                8578 bp    DNA     linear   PAT 10-SEP-2001
DEFINITION  Sequence 17 from Patent WO0161011.
ACCESSION   AX224825
VERSION     AX224825.1  GI:15554917
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 8578)
  AUTHORS   Donoho,G., Scoville,J., Turner,C.A., Friedrich,G., Zambrowicz,B.
            and Sands,A.T.
  TITLE     Novel human thrombospondin repeat proteins and polynucleotides
            encoding the same
  JOURNAL   Patent: WO 0161011-A 17 23-AUG-2001;
            Lexicon Genetics Incorporated (US)
FEATURES             Location/Qualifiers
     source          1..8578
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT     2440 a   1827 c   2076 g   2235 t
ORIGIN

  Query Match             74.0%;  Score 3755;  DB 6;  Length 8578;
  Best Local Similarity   79.7%;  Pred. No. 0;
  Matches 5076;  Conservative    0;  Mismatches    0;  Indels 1291;  Gaps    3;

Qy          1 ATGGCTTCCTGGACGAGCCCCTGGTGGGTGCTGATAGGATGGTCTTCATGCACTCTCCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        179 ATGGCTTCCTGGACGAGCCCCTGGTGGGTGCTGATAGGATGGTCTTCATGCACTCTCCC 238

Qy         61 CTCCCGCAGACCACAGCTGAGAAATCTCCTGGAGCCTATTTCCTTCCCGACTTGGCACTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        239 CTCCCGCAGACCACAGCTGAGAAATCTCCTGGAGCCTATTTCCTTCCCGACTTGGCACTT 298
```

```
Qy    121 TCTCCTCAGGGAAGTTTTCTGGAAGACACAACAGGGGAGCAGTTCCTCACTTATCGCTAT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    299 TCTCCTCAGGGAAGTTTTCTGGAAGACACAACAGGGGAGCAGTTCCTCACTTATCGCTAT 358

Qy    181 GATGACCAGACCTCAAGAAACACTCGTTCAGATGAAGACAAAGATGGCAACTGGGATGCT 240
          ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db    359 GATGACCAGACCTCAAGAAACACTCGTTCAGATGAAGACAAAGATGGCAACTGGGATGCT 418

Qy    241 TGGGGCGACTGGAGTGACTGCTCCCGGACCTGTGGGGGAGGAGCATCATATTCTCTGCGG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    419 TGGGGCGACTGGAGTGACTGCTCCCGGACCTGTGGGGGAGGAGCATCATATTCTCTGCGG 478

Qy    301 AGATGTTTGACTGGAAGGAATTGTGAAGGGCAGAACATTCGGTACAAGACATGCAGCAAT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    479 AGATGTTTGACTGGAAGGAATTGTGAAGGGCAGAACATTCGGTACAAGACATGCAGCAAT 538

Qy    361 CATGACTGCCCTCCAGATGCAGAAGATTTCAGAGCCCAGCAGTGCTCAGCCTACAATGAT 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    539 CATGACTGCCCTCCAGATGCAGAAGATTTCAGAGCCCAGCAGTGCTCAGCCTACAATGAT 598

Qy    421 GTCCAGTATCAGGGGCATTACTATGAATGGCTTCCACGATATAATGATCCTGCTGCCCCG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    599 GTCCAGTATCAGGGGCATTACTATGAATGGCTTCCACGATATAATGATCCTGCTGCCCCG 658

Qy    481 TGTGCACTCAAGTGTCATGCACAAGGACAAAACTTGGTGGTGGAGCTGGCACCTAAGGTA 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    659 TGTGCACTCAAGTGTCATGCACAAGGACAAAACTTGGTGGTGGAGCTGGCACCTAAGGTA 718

Qy    541 CTGGATGGAACTCGTTGCAACACGGACTCCTTGGACATGTGTATCAGTGGCATCTGTCAG 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    719 CTGGATGGAACTCGTTGCAACACGGACTCCTTGGACATGTGTATCAGTGGCATCTGTCAG 778

Qy    601 GCAGTGGGCTGCGATCGGCAACTGGGAAGCAATGCCAAGGAGGACAACTGTGGAGTCTGT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    779 GCAGTGGGCTGCGATCGGCAACTGGGAAGCAATGCCAAGGAGGACAACTGTGGAGTCTGT 838

Qy    661 GCCGGCGATGGCTCCACCTGCAGGCTTGTACGGGGACAATCAAAGTCACACGTTTCTCCT 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    839 GCCGGCGATGGCTCCACCTGCAGGCTTGTACGGGGACAATCAAAGTCACACGTTTCTCCT 898

Qy    721 GAAAAAAGAGAAGAAAATGTAATTGCTGTTCCTTTGGGAAGTCGAAGTGTGAGAATTACA 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    899 GAAAAAAGAGAAGAAAATGTAATTGCTGTTCCTTTGGGAAGTCGAAGTGTGAGAATTACA 958

Qy    781 GTGAAAGGACCTGCCCACCTCTTTATTGAATCAAAAACACTTCAAGGAAGCAAAGGAGAA 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    959 GTGAAAGGACCTGCCCACCTCTTTATTGAATCAAAAACACTTCAAGGAAGCAAAGGAGAA 1018

Qy    841 CACAGCTTTAACAGCCCCGGCGTCTTTGTCGTAGAAAACACAACAGTGGAATTTCAGAGG 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1019 CACAGCTTTAACAGCCCCGGCGTCTTTGTCGTAGAAAACACAACAGTGGAATTTCAGAGG 1078

Qy    901 GGCTCCGAGAGGCAAACTTTTAAGATTCCAGGACCTCTGATGGCTGATTTCATCTTCAAG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1079 GGCTCCGAGAGGCAAACTTTTAAGATTCCAGGACCTCTGATGGCTGATTTCATCTTCAAG 1138

Qy    961 ACCAGGTACACTGCAGCCAAAGACAGCGTGGTTCAGTTCTTCTTTTACCAGCCCATCAGT 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1139 ACCAGGTACACTGCAGCCAAAGACAGCGTGGTTCAGTTCTTCTTTTACCAGCCCATCAGT 1198

Qy   1021 CATCAGTGGAGACAAACTGACTTCTTTCCCTGCACTGTGACGTGTGGAGGAGGTTATCAG 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1199 CATCAGTGGAGACAAACTGACTTCTTTCCCTGCACTGTGACGTGTGGAGGAGGTTATCAG 1258

Qy   1081 CTCAATTCTGCTGAATGTGTGGATATCCGCTTGAAGAGGGTAGTTCCTGACCATTATTGT 1140
          |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db   1259 CTCAATTCTGCTGAATGTGTGGATATCCGCTTGAAGAGGGTAGTTCCTGACCATTATTGT 1318

Qy   1141 CACTACTACCCTGAAAATGTAAAACCAAAACCAAAACTGAAGGAATGCAGCATGGATCCC 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1319 CACTACTACCCTGAAAATGTAAAACCAAAACCAAAACTGAAGGAATGCAGCATGGATCCC 1378

Qy   1201 TGCCCATCAAGTGATGGATTTAAAGAGATAATGCCCTATGACCACTTCCAACCTCTTCCT 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1379 TGCCCATCAAGTGATGGATTTAAAGAGATAATGCCCTATGACCACTTCCAACCTCTTCCT 1438

Qy   1261 C--GCTGGGAACATAATCCTTGGACTGCATGTTCCGTGTCCTGTGGAGGAGGGATTCAGA 1318
          |  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1439 CGAGCTGGGAACATAATCCTTGGACTGCATGTTCCGTGTCCTGTGGAGGAGGGATTCAGA 1498

Qy   1319 GACGGAGCTTTGTGTGTGTAGAGGAATCCATGCATGGAGAGATATTGCAGGTGGAAGAAT 1378
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1499 GACGGAGCTTTGTGTGTGTAGAGGAATCCATGCATGGAGAGATATTGCAGGTGGAAGAAT 1558

Qy   1379 GGAAGTGCATGTACGCACCCAAACCCAAGGTTATGCAAACTTGTAATCTGTTTGATTGCC 1438
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1559 GGAAGTGCATGTACGCACCCAAACCCAAGGTTATGCAAACTTGTAATCTGTTTGATTGCC 1618

Qy   1439 CCAAGTGGATTGCCATGGAGTGGTCTCAGTGCACAGTGACTTGTGGCCGAGGGTTACGGT 1498
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1619 CCAAGTGGATTGCCATGGAGTGGTCTCAGTGCACAGTGACTTGTGGCCGAGGGTTACGGT 1678

Qy   1499 ACCGGGTTGTTCTGTGTATTAACCACCGCGGAGAGCATGTTGGGGGCTGCAATCCACAAC 1558
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1679 ACCGGGTTGTTCTGTGTATTAACCACCGCGGACAGCATGTTGGGGGCTGCAATCCACAAC 1738

Qy   1559 TGAAGTTACACATCAAAGAAGAATGTGTCATTCCCATCCCGTGTTATAAACCAAAAGAAA 1618
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1739 TGAAGTTACACATCAAAGAAGAATGTGTCATTCCCATCCCGTGTTATAAACCAAAAGAAA 1798

Qy   1619 AAAGTCCAGTGGAAGCAAAATTGCCTTGGCTGAAACAAGCACAAGAACTAGAAGAGACCA 1678
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1799 AAAGTCCAGTGGAAGCAAAATTGCCTTGGCTGAAACAAGCACAAGAACTAGAAGAGACCA 1858

Qy   1679 GAATAGCAACAGAAGAACCAACGTTCATTCCAGAACCCTGGTCAGCCTGCAGTACCACGT 1738
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1859 GAATAGCAACAGAAGAACCAACGTTCATTCCAGAACCCTGGTCAGCCTGCAGTACCACGT 1918

Qy   1739 GTGGGCCAGGTGTGCAGGTCCGCCAGGTGAAGTGCCGTGTGCTCCTCACATTCACGCAGA 1798
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1919 GTGGCCCAGGTGTGCAGGTCCGCGAGGTGAAGTGCCGTGTGCTCCTCACATTCACGCAGA 1978

Qy   1799 CTGAGACTGAGCTGCCCGAGGAAGAGTGTGAAGCCCCCAAGCTGCCCACCGAACGGCCCT 1858
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1979 CTGAGACTGAGCTGCCCGAGGAAGAGTGTGAAGCCCCCAAGCTGCCCACCGAACGGCCCT 2038

Qy   1859 GCCTCCTGGAAGCATGTGATGAGAGCCCGGCCTCCCGAGAGCTAGACATCCCTCTCCCTG 1918
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2039 GCCTCCTGGAAGCATGTGATGAGAGCCCGGCCTCCCGAGAGCTAGACATCCCTCTCCCTG 2098

Qy   1919 AGGACAGTGAGACGACTTACGACTGGGAGTACGCTGGGTTCACCCCTTGCACAGCAACAT 1978
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2099 AGGACAGTGAGACGACTTACGACTGGGAGTACGCTGGGTTCACCCCTTGCACAGCAACAT 2158

Qy   1979 GCTTGGCAGGCCATCAAGAAGCCATAGCAGTGTGCTTACATATCCAGACCCAGCAGACAG 2038
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2159 GCTTGGCAGGCCATCAAGAAGCCATAGCAGTGTGCTTACATATCCAGACCCAGCAGACAG 2218

Qy   2039 TCAATGACAGCTTGTGTGATATGGTCCACCGTCCTCCAGCCATGAGCCAGGCCTGTAACA 2098
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2219 TCAATGACAGCTTGTGTGATATGGTCCACCGTCCTCCAGCCATGAGCCAGGCCTGTAACA 2278

Qy   2099 CAGAGCCCTGTCCCCCCAGG---------------------------------------- 2118
          ||||||||||||||||||||
Db   2279 CAGAGCCCTGTCCCCCCAGGACAGAGCCCAGCAGCTTGTAGAAGCATGCCGGGTTACATAA 2338

Qy   2119 ------------------------------------------------------------ 2118

Db   2339 TGGTCCTGCTAGTCTGAGGAGAGCCTTCTTCTCTAACAGGATTCAACACTGCTAGGGAAG 2398

Qy   2119 ------------------------------------------------------------ 2118

Db   2399 AAAGGAGGAAAGCAAGAGGCAATAGTGATGTGTTTCTGTACCAGCTTGTTACCTATTTCT 2458

Qy   2119 ------------------------------------------------------------ 2118
```

```
Db     2459 TGATATAAAAAACAATTCTTAATTGATTCATTCTTTTGAATAAGAAATTTCCTCTCATT 2518

Qy     2121                                                TGGCATGGGCTTT 2133
                                                           ||||| || | |
Db     2519 TTCTTAAATAAAAACAGCTCTATTTCCAAAAAAAAAAAAAAAAAATGGCATGGGCTTT 2578

Qy     2134 TGGGGGCCTGCTCAGCTACTTGTGGAGTTGGAATTCAGACCGAGATCGTAACTGCCTG 2193
            ||||||||||||||||||   |||||||||||||||||||| ||||| | ||| |  |
Db     2579 TGGGGGCCTGCTCAGCTACCTGTGGAGTTGGAATTCAGACCGAGATGTGACTGCCTG 2638

Qy     2194 CACCCAGGGAGACCCTCCCCTCCTGAGGAGTGCCAGATGAAAAGCCCCATGCTTTA 2253
            ||||||||||||||||||||||||||||||||||||||||||||||:||||||||| |
Db     2639 CACCCAGGGAGACCCTGCCCTCCTGAGGAGTGCCAGATGAAAAGCCCCATGCTTTA 2698

Qy     2254 CAAGCATGCAATCAGTTTGACTGCCCTCCTGGCTGGCACATTGAAGAATGGCAGCAGTGT 2313
            ||||||||||||||||||||  |||||||||||||||||||||||||| ||||  | |
Db     2699 CAAGCATGCAATCAGTTTGACTGCCCTCCTGGCTGGCACATTGAAGAATGGCAGCAGTGT 2758

Qy     2314 TCCAGGACTTGTGGGGGGGGAACTCAGAAGAGAAGAGTCACCTGCTGGCACCTGCTAACT 2373
            |||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db     2759 TCCAGGACTTGTGGGGGGGGAACTCAGAAGAGAAGAGTCACCTGCTGGCACCTGCTAACT 2818

Qy     2374 GATGGCAGCTTTTTGAATCTCTCAGATGAATTGTGCCAAGGACCCAAGGCATCGTCTCAC 2433
            |||||||||||| |||||||| ||||||||||||||||||||||| |||||||||  |
Db     2819 GATGGCAGCTTTTTGAATCTCTCAGATGAATTGTGCCAAGGACCCAAGGCATCGTCTCAC 2878

Qy     2434 AAGTCCTGTGCCAGGACAGACTGTCCTCCACATTTAGCTGTGGGAGACTGGCTGAAG- 2490
            ||||||||||||||||  ||||||||||||||||||||||||||||||     |||||
Db     2879 AAGTCCTGTGCCAGGACAGACTGTCCTCCACATTTAGCTGTGGGAGACTGGCTGAAGGAT 2938

Qy     2491 ------------------------------------------------------------ 2490

Db     2939 CATTCAATGCAAGAGGACAATGGAGCAGGATCTACACAATTCTAAAGAAAAGTAAGCATG 2998

Qy     2491 ------------------------------------------------------------ 2490

Db     2999 ACTCAAGGATTTCCTCTTCATCCTGTGTTCTTCATGTAGAGAGACAGCAGAGAGGCAGTC 3058

Qy     2491 ------------------------------------------------------------ 2490

Db     3059 AGAGAATACTGTCTGATAAGCCCTTGAAAAAGCTGTAGGGCCAAGATGAGATACAGAGAT 3118

Qy     2491 ------------------------------------------------------------ 2490

Db     3119 GACTCAAAACAGAGAATCCAGGAATGCATACATCCTGGTAAAAAAGGTGGGAGATGAGTA 3178

Qy     2491 ------------------------------------------------------------ 2490

Db     3179 ATAAATTCATTTGTGTAGGATTAAGACTAATCAACTAACAATTATATTATAGAACATAAC 3238

Qy     2491 ------------------------------------------------------------ 2490

Db     3239 ATAAATATTAGAAATCTTGAATCATCTAAATAATAAAATGAAAACTAATTGAGATTTT 3298

Qy     2491 ------------------------------------------------------------ 2490

Db     3299 AGAGATGAGGTAGATGATATATTTTGGCTGTGTCCCACCCAAATCTCATCTTGAATTGT 3358

Qy     2491 ------------------------------------------------------------ 2490

Db     3359 AGTTCCCATAATTCCCATGTGTTGTGGGAGGGACCTAGTTGGAGATAATTGAATCATTCC 3418

Qy     2491 ------------------------------------------------------------ 2490

Db     3419 GGCATTTTCCCTCATATCGTTCTCTGGTGGTAAATGAGTCTCACGAGATCTGATGGTTT 3478

Qy     2491 ------------------------------------------------------------ 2490

Db     3479 TATAAGGGGTTTCCCTTTTCGCTTCGGCTCTCATTCTCTCTTGCCTGCTGCCATGTAAGAC 3538

Qy     2491 ------------------------------------------------------------ 2490

Db     3539 GTCCCTTTGCCCTTCCTTTGTCTCTGCCATGATTGTGAGGCTTCCCCTATCCACTGGAA 3598

Qy     2491 ------------------------------------------------------------ 2490

Db     3599 CTGAGTCTATTAAACCTCTTTCCTTTATAACTTACCCAGTCTTGGGTATGTCTTTATTAA 3658

Qy     2491 ------------------------------------------------------------ 2490

Db     3659 CAACATAAGATTGGACTAATACAGTAGAGGAAATGTAAGCTTGCTTATTTCCTCATCCTT 3718

Qy     2491 ------------------------------------------------------------ 2490

Db     3719 CTTAGTAGCAAGTCAATAAATACTGTCCTAAGTCAAATTTCTATTAAAAATAACTATTCA 3778

Qy     2491 ------------------------------------------------------------ 2490

Db     3779 AATCTCTTGTTGGTTTATTTAATCTTCTTTATTAACTTAGAGTGTTCTTTCGGGAATTA 3838

Qy     2491 ------------------------------------------------------------ 2490

Db     3839 ATCATGGTTTAAAAAATATCAAACATTCAACAACTCTAATTTTACTTTAATGTCTTTTT 3898

Qy     2491 ------------------------------------------------------------ 2490

Db     3899 TCTAATATATCTAATAAAATTGCATTAAATTTTTAAGTTGAAAAAAAAAAAAAAAAAAAA 3958

Qy     2491 -TGTTCTGTCAGTTGTGGTGTTGGAATCCAGAGAAGAAAGCAGGTGTGTCAAAGGCTGGC 2549
             |||||||||||| | ||||| |||||||||||||||||| | ||||||||||||||||
Db     3959 ATGTTCTGTCAGTTGTGGTGTTGGAATCCAGAGAAGAAAGCAGGTGTGTCAAAGGCTGGC 4018

Qy     2550 AGCCAAAGGTCGGCGCATCCCCCTCAGTGAGATGATGTGCAGGGATCTACCACGGTTCCC 2609
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4019 AGCCAAAGGTCGGCGCATCCCCCTCAGTGAGATGATGTGCAGGGATCTACCAGGGTTCCC 4078

Qy     2610 TCTTGTAAGATCTTGCCAGATGCCTGAGTGCAGTAAAATCAAATCAGAGATGAAGACAAA 2669
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4079 TCTTGTAAGATCTTGCCAGATGCCTGAGTGCAGTAAAATCAAATCAGAGATGAAGACAAA 4138

Qy     2670 ACTTGGTGAGCAGGGTCCGCAGATCCTCAGTGTCCAGAGAGTCTACATTCAGACAAGGGA 2729
            |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db     4139 ACTTGGTGAGCAGGGTCCGCAGATCCTCAGTGTCCAGAGAGTCTACATTCAGACAAGGGA 4198

Qy     2730 AGAGAAGCGTATTAACCTGACCATTGGTAGCAGAGCCTATTTGCTGCCCAACACATCCGT 2789
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4199 AGAGAAGCGTATTAACCTGACCATTGGTAGCAGAGCCTATTTGCTGCCCAACACATCCGT 4258

Qy     2790 GATTATTAAGTGCCCCGTGCGACGATTCCAGAAATCTCTGATCCAGTGGGAGAAGGATGG 2849
            ||||||||||||||||||||||||||||||||||||||||     |||||| ||||||
Db     4259 GATTATTAAGTGCCCCGTGCGACGATTCCAGAAATCTCTGATCCAGTGGGAGAAGGATGG 4318

Qy     2850 CCGTTGCCTGCAGAACTCCAAACGGCTTGGCATCACCAAGTCAGGCTCACTAAAAATCCA 2909
            |||||||||||||||| | |||||||||||||||||||||||||||||||| ||||||
Db     4319 CCGTTGCCTGCAGAACTCCAAACGGCTTGGCATCACCAAGTCAGGCTCACTAAAAATCCA 4378

Qy     2910 CGGTCTTGCTGCCCCGACATCGGCGTGTACCGGTGCATCGCAGCCTCTGCACAGGAAAC 2969
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4379 CGGTCTTGCTGCCCCGACATCGGCGTGTACCGGTGCATCGCAGCCTCTGCACAGGAAAC 4438

Qy     2970 AGTTGTGCTCAAGCTCATTGGTACTGACAACCGGCTCATCGCAGGCCCAGCCCTCAGGGA 3029
            |||||||||||||| |||||||||||||||||||||||| |||||||| ||||| ||||
Db     4439 AGTTGTGCTCAAGCTCATTGGTACTGACAACCGGCTCATCGCAGGCCCAGCCCTCAGGGA 4498

Qy     3030 GCCTATGAGGGAATATCCTGGGATGGACCACAGCGAAGCAATAGTTGGGAGTCACATG 3089
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4499 GCCTATGAGGGAATATCCTGGGATGGACCACAGCGAAGCAATAGTTGGGAGTCACATG 4558

Qy     3090 GCACAAAATGAGGCAAATGTGGAATAACAAAAATGACCTTTATCTGGATGATGACCACAT 3149
            |||||||||||||||||||||||||||||||||| |||  ||     ||||||||| | ||
Db     4559 GCACAAAATGAGGCAAATGTGGAATAACAAAAATGACCTTTATCTGGATGATGACCACAT 4618

Qy     3150 TAGTAACCAGCCTTTCTTGAGAGCTCTGTTAGGCCACTGCAGCAATTCTGCAGGAAGCAC 3209
             |||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db     4619 TAGTAACCAGCCTTTCTTGAGAGCTCTGTTAGGCCACTGCAGCAATTCTGCAGGAAGCAC 4678
```

```
Qy       3210 CAACTCCTGGGAGTTGAAGAATAAGCAGTTTGAAGCAGCAGTTAAACAAGGAGCATATAG 3269
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4679 CAACTCCTGGGAGTTGAAGAATAAGCAGTTTGAAGCAGCAGTTAAACAAGGAGCATATAG 4738

Qy       3270 CATGGATACAGCCCAGTTTGATGAGCTGATAAGAAACATGAGTCAGCTCATGGAAACCGG 3329
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4739 CATGGATACAGCCCAGTTTGATGAGCTGATAAGAAACATGAGTCAGCTCATGGAAACCGG 4798

Qy       3330 AGAGGTCAGCGATGATCTTGCGTCCCAGCTGATATATCAGCTGCTGGCCGAATTAGCCAA 3389
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4799 AGAGGTCAGCGATGATCTTGCGTCCCAGCTGATATATCAGCTGCTGGCCGAATTAGCCAA 4858

Qy       3390 GGCACAGCCAACACACATGCAGTGGCGGGGCATCCAGGAAGAGACACCTCCTGCTGCTCA 3449
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4859 GGCACAGCCAACACACATGCAGTGGCGGGGCATCCAGGAAGAGACACCTCCTGCTGCTCA 4918

Qy       3450 GCTCAGAGGGGAAACAGGGAGTGTGTCCCAAAGCTCGCATGCAAAAAACTCAGGCAAGCT 3509
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4919 GCTCAGAGGGGAAACAGGGAGTGTGTCCCAAAGCTCGCATGCAAAAAACTCAGGCAAGCT 4978

Qy       3510 GACATTCAAGCCGAAAGGACCTGTTCTCATGAGGCAAAGCCAACCTCCCTCAATTTCATT 3569
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4979 GACATTCAAGCCGAAAGGACCTGTTCTCATGAGGCAAAGCCAACCTCCCTCAATTTCATT 5038

Qy       3570 TAATAAAACAATAAATTCCAGGATTGGAAATACAGTATACATTACAAAAAGGACAGAGGT 3629
              ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db       5039 TAATAAAACAATAAATTCCAGGATTGGAAATACAGTATACATTACAAAAAGGACAGAGGT 5098

Qy       3630 CATCAATATACTGTGTGACCTTATTACCCCCAGTGAGGCCACATATACATGGACCAAGGA 3689
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5099 CATCAATATACTGTGTGACCTTATTACCCCCAGTGAGGCCACATATACATGGACCAAGGA 5158

Qy       3690 TGGAACCTTGTTACAGCCCTCAGTAAAAATAATTTTGGATGGAACTGGGAAGATACAGAT 3749
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5159 TGGAACCTTGTTACAGCCCTCAGTAAAAATAATTTTGGATGGAACTGGGAAGATACAGAT 5218

Qy       3750 ACAGAATCCTACAAGGAAAGAACAAGGCATATATGAATGTTCTGTAGCTAATCATCTTGG 3809
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5219 ACAGAATCCTACAAGGAAAGAACAAGGCATATATGAATGTTCTGTAGCTAATCATCTTGG 5278

Qy       3810 TTCAGATGTGGAAAGTTCTTCTGTGCTGTATGCAGAGGCACCTGTCATCTTGTCTGTTGA 3869
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5279 TTCAGATGTGGAAAGTTCTTCTGTGCTGTATGCAGAGGCACCTGTCATCTTGTCTGTTGA 5338

Qy       3870 AAGAAATATCACCAAACCAGAGCACAACCATCTGTCTGTTGTGGTTGGAGGCATCGTGGA 3929
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5339 AAGAAATATCACCAAACCAGAGCACAACCATCTGTCTGTTGTGGTTGGAGGCATCGTGGA 5398

Qy       3930 GGCAGCCCTTGGAGCAAACGTGACAATCCGATGTCCTGTAAAAGGTGTCCCTCAGCCTAA 3989
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5399 GGCAGCCCTTGGAGCAAACGTGACAATCCGATGTCCTGTAAAAGGTGTCCCTCAGCCTAA 5458

Qy       3990 TATAACTTGGTTGAAGAGAGGAGGATCTCTGAGTGGCAATGTTTCCTTGCTTTTCAATGG 4049
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5459 TATAACTTGGTTGAAGAGAGGAGGATCTCTGAGTGGCAATGTTTCCTTGCTTTTCAATGG 5518

Qy       4050 ATCCCTGTTGTTGCAGAATGTTTCCCTTGAAAATGAAGGAACCTACGTCTGCATAGCCAC 4109
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5519 ATCCCTGTTGTTGCAGAATGTTTCCCTTGAAAATGAAGGAACCTACGTCTGCATAGCCAC 5578

Qy       4110 CAATGCTCTTGGAAAGGCAGTGGCAACATCTGTACTCCACTTGCTGGAACGAAGATGGCC 4169
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5579 CAATGCTCTTGGAAAGGCAGTGGCAACATCTGTACTCCACTTGCTGGAACGAAGATGGCC 5638

Qy       4170 AGAGAGTAGAATCGTATTTCTGCAAGGACATAAAAAGTACATTCTCCAGGCAACCAACAC 4229
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5639 AGAGAGTAGAATCGTATTTCTGCAAGGACATAAAAAGTACATTCTCCAGGCAACCAACAC 5698

Qy       4230 TAGAACCAACAGCAATGACCCAACAGGAGAACCCCCGCCTCAAGAGCCTTTTTGGCAGCC 4289
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5699 TAGAACCAACAGCAATGACCCAACAGGAGAACCCCCGCCTCAAGAGCCTTTTTGGCAGCC 5758

Qy       4290 TGGTAACTGGTCACATTGTTCTGCCACCTGTGGTCATTTGGGAGCCCGCATTCAGAGACC 4349
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5759 TGGTAACTGGTCACATTGTTCTGCCACCTGTGGTCATTTGGGAGCCCGCATTCAGAGACC 5818

Qy       4350 CCAGTGTGTGATGGCCAATGGGCAGGAAGTGAGTGAGGCCCTGTGTGATCACCTCCAGAA 4409
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5819 CCAGTGTGTGATGGCCAATGGGCAGGAAGTGAGTGAGGCCCTGTGTGATCACCTCCAGAA 5878

Qy       4410 GCCACTGGCTGGGTTTGAGCCCTGTAACATCCGGGACTGCCCAGCGAGGTGGTTCACAAG 4469
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5879 GCCACTGGCTGGGTTTGAGCCCTGTAACATCCGGGACTGCCCAGCGAGGTGGTTCACAAG 5938

Qy       4470 TGTGTGGTCACAGTGCTCTGTGTCTTGCGGTGAAGGATACCACAGTCGGCAGGTGACGTG 4529
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5939 TGTGTGGTCACAGTGCTCTGTGTCTTGCGGTGAAGGATACCACAGTCGGCAGGTGACGTG 5998

Qy       4530 CAAGCGGACAAAAGCCAATGGAACTGTGCAGGTGGTGTCTCCAAGAGCATGTGCCCCTAA 4589
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5999 CAAGCGGACAAAAGCCAATGGAACTGTGCAGGTGGTGTCTCCAAGAGCATGTGCCCCTAA 6058

Qy       4590 AGACCGGCCTCTGGGAAGAAAACCATGTTTTTGGTCATCCATGTGTTCAGTGGAACCAGG 4649
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6059 AGACCGGCCTCTGGGAAGAAAACCATGTTTTTGGTCATCCATGTGTTCAGTGGAACCAGG 6118

Qy       4650 GAACCGGTGTCCTGGACGTTGCATGGGCCGTGCTGTGAGGATGCAGCAGCGTCACACAGC 4709
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6119 GAACCGGTGTCCTGGACGTTGCATGGGCCGTGCTGTGAGGATGCAGCAGCGTCACACAGC 6178

Qy       4710 TTGTCAACACAACAGCTCTGACTCCAACTGTGATGACAGAAAGAGACCCACCTTAAGAAG 4769
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6179 TTGTCAACACAACAGCTCTGACTCCAACTGTGATGACAGAAAGAGACCCACCTTAAGAAG 6238

Qy       4770 GAACTGCACATCAGGGGCCTGTGATGTGTGTTGGCACACAGGCCCTTGGAAGCCCTGTAC 4829
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6239 GAACTGCACATCAGGGGCCTGTGATGTGTGTTGGCACACAGGCCCTTGGAAGCCCTGTAC 6298

Qy       4830 AGCAGCCTGTGGCAGGGGTTTCCAGTCTCGGAAAGTCGACTGTATCCACACAAGGAGTTG 4889
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6299 AGCAGCCTGTGGCAGGGGTTTCCAGTCTCGGAAAGTCGACTGTATCCACACAAGGAGTTG 6358

Qy       4890 CAAACCTGTGGCCAAGAGACACTGTGTACAGAAAAAGAAACCAATTTCCTGGCGGCACTG 4949
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6359 CAAACCTGTGGCCAAGAGACACTGTGTACAGAAAAAGAAACCAATTTCCTGGCGGCACTG 6418

Qy       4950 TCTTGGGCCCTCCTGTGATAGAGACTGCACAGACACAACTCACTACTGTATGTTTGTAAA 5009
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6419 TCTTGGGCCCTCCTGTGATAGAGACTGCACAGACACAACTCACTACTGTATGTTTGTAAA 6478

Qy       5010 ACATCTTAATTTGTGTTCTCTAGACCGCTACAAACAAAGGTGCTGCCAGTCATGTCAAGA 5069
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6479 ACATCTTAATTTGTGTTCTCTAGACCGCTACAAACAAAGGTGCTGCCAGTCATGTCAAGA 6538

Qy       5070 GGGATAA 5076
              |||||||
Db       6539 GGGATAA 6545


RESULT 7
AX235377
LOCUS       AX235377                5107 bp    DNA     linear   PAT 11-SEP-2001
DEFINITION  Sequence 24 from Patent WO0162928.
ACCESSION   AX235377
VERSION     AX235377.1  GI:15593909
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 5107)
  AUTHORS   Vernet,C.A., Fernandes,E., Shimkets,R.A., Macdougall,J. and
            Spaderna,S.K.
  TITLE     Polypeptides and nucleic acids encoding same
  JOURNAL   Patent: WO 0162928-A 24 30-AUG-2001;
            Curagen Corporation (US)
```

```
FEATURES             Location/Qualifiers
     source          1..5107
                     /organism "Homo sapiens"
                     /db_xref "taxon:9606"
BASE COUNT     1444 a   1262 c   1257 g   1144 t
ORIGIN

  Query Match             60.5%;  Score 3072.4;  DB 6;  Length 5107;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 3073;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2003 TAGCAGTGTGCTTACATATCCAGACCCAGCAGACAGTCAATGACAGCTTGTGTGATATGG 2062
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TAGCAGTGTGCTTACATATCCAGACCCAGCAGACAGTCAATGACAGCTTGTGTGATATGG 60

Qy       2063 TCCACCGTCCTCCAGCCATGACCCAGGCCTGTAACACAGAGCCCTGTCCCCCCAGGTGGC 2122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCCACCGTCCTCCAGCCATGACCCAGGCCTGTAACACAGAGCCCTGTCCCCCCAGGTGGC 120

Qy       2123 ATGTGGGCTCTTGGGGGCCCTGCTCAGCTACCTGTGGAGTTGGAATTCAGACCCGAGATG 2182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ATGTGGGCTCTTGGGGGCCCTGCTCAGCTACCTGTGGAGTTGGAATTCAGACCCGAGATG 180

Qy       2183 TGTACTGCCTGCACCCAGGGGAGACCCCTGCCCCTCCTGAGGAGTGCCGAGATGAAAAGC 2242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGTACTGCCTGCACCCAGGGGAGACCCCTGCCCCTCCTGAGGAGTGCCGAGATGAAAAGC 240

Qy       2243 CCCATGCTTTACAAGCATGCAATCAGTTTGACTGCCCTCCTGGCTGGCACATTGAAGAAT 2302
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCCATGCTTTACAAGCATGCAATCAGTTTGACTGCCCTCCTGGCTGGCACATTGAAGAAT 300

Qy       2303 GGCAGCAGTGTTCCAGGACTTGTGGCGGGGGAACTCAGAACAGAAGAGTCACCTGTCGGC 2362
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGCAGCAGTGTTCCAGGACTTGTGGCGGGGGAACTCAGAACAGAAGAGTCACCTGTCGGC 360

Qy       2363 AGCTGCTAACGGATGGCAGCTTTTTGAATCTCTCAGATGAATTGTGCCAAGGACCTAACC 2422
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AGCTGCTAACGGATGGCAGCTTTTTGAATCTCTCAGATGAATTGTGCCAAGGACCTAACC 420

Qy       2423 CATCGTCTCACAAGTCCTGTGCCAGGACAGACTGTCCTCCACATTTAGCTGTGGGAGACT 2482
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CATCGTCTCACAAGTCCTGTGCCAGGACAGACTGTCCTCCACATTTAGCTGTGGGAGACT 480

Qy       2483 GGTCGAAGTGTTCTGTCAGTTGTGGTGTTGGAATCCAGAGAAGAAAGCAGGTGTGTCAAA 2542
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GGTCGAAGTGTTCTGTCAGTTGTGGTGTTGGAATCCAGAGAAGAAAGCAGGTGTGTCAAA 540

Qy       2543 GGCTGGCAGCCAAAGTGTCGGGCATCCCCCTCAGTGAGATGATGTGCAGGGATCTACCAG 2602
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GGCTGGCAGCCAAAGTGTCGGGCATCCCCCTCAGTGAGATGATGTGCAGGGATCTACCAG 600

Qy       2603 GGTTCCCTCTTGTAAGATCTTGCCAGATGCCTGAGTGCAGTAAAATCAAATCAGAGATGA 2662
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GGTTCCCTCTTGTAAGATCTTGCCAGATGCCTGAGTGCAGTAAAATCAAATCAGAGATGA 660

Qy       2663 AGACAAAACTTGGTGAGCAGGGTCCGCAGATCCTCAGTGTCCAGAGAGTCTACATTCAGA 2722
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AGACAAAACTTGGTGAGCAGGGTCCGCAGATCCTCAGTGTCCAGAGAGTCTACATTCAGA 720

Qy       2723 CAAGGGAAGAGAAGCGTATTAACCTGACCATTGGTAGCAGAGCCTATTTGCTGCCCAACA 2782
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CAAGGGAAGAGAAGCGTATTAACCTGACCATTGGTAGCAGAGCCTATTTGCTGCCCAACA 780

Qy       2783 CATCCGTGATTATTAAGTGCCCCGTGCGACGATTCCAGAAATCTCTGATCCAGTGGGAGA 2842
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CATCCGTGATTATTAAGTGCCCCGTGCGACGATTCCAGAAATCTCTGATCCAGTGGGAGA 840

Qy       2843 AGGATGGCCGTTGGCTGCAGAACTCCAAACGGCTTGGCATCACCAAGTCAGGCTCACTAA 2902
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AGGATGGCCGTTGGCTGCAGAACTCCAAACGGCTTGGCATCACCAAGTCAGGCTCACTAA 900

Qy       2903 AAATCCACGGTCTTGCTGCCCCCGACATCGGCGTGTACCGGTGCATTGCAGGCTCTGCAC 2962
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AAATCCACGGTCTTGCTGCCCCCGACATCGGCGTGTACCGGTGCATTGCAGGCTCTGCAC 960

Qy       2963 AGGAAACAGTTGTGCTCAAGCTCATTGGTACTGACAACCGGCTCATCCTATCCTCAGCCC 3022
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 AGGAAACAGTTGTGCTCAAGCTCATTGGTACTGACAACCGGCTCATCCTATCCTCAGCCC 1020

Qy       3023 TCAGGGAGCCTATGAGGGAATATCCTGGGATGGACCACAGCGAAGCCAATAGTTTGGGAG 3082
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TCAGGGAGCCTATGAGGGAATATCCTGGGATGGACCACAGCGAAGCCAATAGTTTGGGAG 1080

Qy       3083 TCACATGGCACAAAATGAGGCAAATGTGGAATAACAAAAATGACCTTTATCTGGATGATG 3142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TCACATGGCACAAAATGAGGCAAATGTGGAATAACAAAAATGACCTTTATCTGGATGATG 1140

Qy       3143 ACCACATTAGTAACCAGCCTTTCTTGAGAGCTCTGTTAGCCCACTGCAGCAACTCTGCAG 3202
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 ACCACATTAGTAACCAGCCTTTCTTGAGAGCTCTGTTAGCCCACTGCAGCAACTCTGCAG 1200

Qy       3203 GAAGCACCAACTCCTGGGAGTTGAAGAATAAGCAGTTTGAAGCAGCAGTTAAACAAGGAG 3262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GAAGCACCAACTCCTGGGAGTTGAAGAATAAGCAGTTTGAAGCAGCAGTTAAACAAGGAG 1260

Qy       3263 CATATAGCATGGATACAGCCCAGTTTGATGAGCTGATAAGAAACATGAGTCAGCTCATGG 3322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CATATAGCATGGATACAGCCCAGTTTGATGAGCTGATAAGAAACATGAGTCAGCTCATGG 1320

Qy       3323 AAACCGGAGAGGTCAGCGATGATCTTGCGTCCCAGCTGATATATCAGCTGGTGGCCGAAT 3382
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 AAACCGGAGAGGTCAGCGATGATCTTGCGTCCCAGCTGATATATCAGCTGGTGGCCGAAT 1380

Qy       3383 TAGCCAAGGCACAGCCAACACACATGCAGTGGCGGGGCATCCAGGAAGAGACACCTCCTG 3442
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 TAGCCAAGGCACAGCCAACACACATGCAGTGGCGGGGCATCCAGGAAGAGACACCTCCTG 1440

Qy       3443 CTGCTCAGCTCAGAGGGAAACAGGGAGTGTGTCCCAAAGCTCCCATGCAAAAAAACTCAG 3502
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 CTGCTCAGCTCAGAGGGAAACAGGGAGTGTGTCCCAAAGCTCCCATGCAAAAAAACTCAG 1500

Qy       3503 GCAAGCTGACATTCAAGCCGAAAGGACCTGTTCTTCATGAGGAAAGCAACCTCCCCTCAA 3562
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 GCAAGCTGACATTCAAGCCGAAAGGACCTGTTCTTCATGAGGAAAGCAACCTCCCCTCAA 1560

Qy       3563 TTTCATTTAATAAAACAATAAATTCCAGGATTGGAAATACAGTATACATTACAAAAAGGA 3622
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 TTTCATTTAATAAAACAATAAATTCCAGGATTGGAAATACAGTATACATTACAAAAAGGA 1620

Qy       3623 CAGAGGTCATCAATATACTGTGTGACCTTATTACCCCCAGTGAGGCCACATATACATGGA 3682
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 CAGAGGTCATCAATATACTGTGTGACCTTATTACCCCCAGTGAGGCCACATATACATGGA 1680

Qy       3683 CCAAGGATGGAACCTTGTTACAGCCCCTCAGTAAAAATAATTTGGATGGAACTGGGAAGA 3742
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 CCAAGGATGGAACCTTGTTACAGCCCCTCAGTAAAAATAATTTGGATGGAACTGGGAAGA 1740

Qy       3743 TACAGATACAGAATCCTACAAGGAAAGAACAAGGCATATATGAATGTTCTGTAGCTAATC 3802
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 TACAGATACAGAATCCTACAAGGAAAGAACAAGGCATATATGAATGTTCTGTAGCTAATC 1800

Qy       3803 ATCTTGGTTCAGATGTGGAAAGCTCTTCTGGGCTGTATGGAGAGGCACCTGTCATCTTGT 3862
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 ATCTTGGTTCAGATGTGGAAAGCTCTTCTGGGCTGTATGGAGAGGCACCTGTCATCTTGT 1860

Qy       3863 CTGTTGAAAGAAATATCACCAAACCAGAGCACAACCATCTGTCTGTTGTGGTTGGAGGCA 3922
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 CTGTTGAAAGAAATATCACCAAACCAGAGCACAACCATCTGTCTGTTGTGGTTGGAGGCA 1920

Qy       3923 TCGTGGAGGCAGCCCTTGGAGCAAACGTGACAATCCGATGTCCTGTAAAAGGTGTCCCTC 3982
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 TCGTGGAGGCAGCCCTTGGAGCAAACGTGACAATCCGATGTCCTGTAAAAGGTGTCCCTC 1980

Qy       3983 AGCTAATATAACTTGGTTGAAGAGAGGAGGATCTCTGAGTGGCAATGTTTTCCTTGCTTT 4042
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 AGCTAATATAACTTGGTTGAAGAGAGGAGGATCTCTGAGTGGCAATGTTTTCCTTGCTTT 2040
```

```
Qy     4043 TCAATGGATCCCTGTTGTTGCAGAATGTTTCCCTTGAAAATGAAGGAACCTACGTCTGCA 4102
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2041 TCAATGGATCCCTGTTGTTGCAGAATGTTTCCCTTGAAAATGAAGGAACCTACGTCTGCA 2100

Qy     4103 TAGCCACCAATGCTCTTGGAAAGGCAGTGGCAACATCTGTACTCCACTTGCTGGAACCAA 4162
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2101 TAGCCACCAATGCTCTTGGAAAGGCAGTGGCAACATCTGTACTCCACTTGCTGGAACCAA 2160

Qy     4163 GATGGCCAGAGAGTAGAATCGTATTTCTGCAAGGACATAAAAAGTACATTCTCCAGGCAA 4222
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2161 GATGGCCAGAGAGTAGAATCGTATTTCTGCAAGGACATAAAAAGTACATTCTCCAGGCAA 2220

Qy     4223 CCAACACTAGAACCAACAGCAATGACCCAACAGGAGAACCCCCGCCTCAAGAGCCTTTTT 4282
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2221 CCAACACTAGAACCAACAGCAATGACCCAACAGGAGAACCCCCGCCTCAAGAGCCTTTTT 2280

Qy     4283 GGGAGCCTGGTAACTGGTCACATTGTTCTGCCACCTGTGGTCATTTGGGAGCCCGCATTC 4342
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2281 GGGAGCCTGGTAACTGGTCACATTGTTCTGCCACCTGTGGTCATTTGGGAGCCCGCATTC 2340

Qy     4343 AGAGACCCCAGTGTGTGATGGCCAATGGGCAGGAAGTGAGTGAGGCCCTGTGTGATCACC 4402
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2341 AGAGACCCCAGTGTGTGATGGCCAATGGGCAGGAAGTGAGTGAGGCCCTGTGTGATCACC 2400

Qy     4403 TCCAGAAGCCACTGGCTGGGTTTGAGCCCTGTAACATCCGGGACTGCCCAGCCAGCTGGT 4462
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2401 TCCAGAAGCCACTGGCTGGGTTTGAGCCCTGTAACATCCGGGACTGCCCAGCCAGCTGGT 2460

Qy     4463 TCACAAGTGTGTGGTCACAGTGCTCTGTGTCTTGCGGTCAAGGATACCACAGTCGGCAGG 4522
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2461 TCACAAGTGTGTGGTCACAGTGCTCTGTGTCTTGCGGTCAAGGATACCACAGTCGGCAGG 2520

Qy     4523 TGACGTGCAAGCGGACAAAAGCCAATGGAACTGTGCAGGTGGTGTCTCCAAGAGCATGTG 4582
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2521 TGACGTGCAAGCGGACAAAAGCCAATGGAACTGTGCAGGTGGTGTCTCCAAGAGCATGTG 2580

Qy     4583 CCCCTAAAGACCGGCCTCTGGGAAGAAAACCATGTTTTGGTCATCCATGTGTTCAGTGGG 4642
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2581 CCCCTAAAGACCGGCCTCTGGGAAGAAAACCATGTTTTGGTCATCCATGTGTTCAGTGGG 2640

Qy     4643 AACCAGGGAACCGGTGTCCTGGACGTTGCATGGGCCGTGCTGTGAGGATGCAGCAGCGTC 4702
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2641 AACCAGGGAACCGGTGTCCTGGACGTTGCATGGGCCGTGCTGTGAGGATGCAGCAGCGTC 2700

Qy     4703 ACACAGCTTGTCAACACAACAGCTCTGACTCCAACTGTGATGACAGAAAGAGACCCACCT 4762
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2701 ACACAGCTTGTCAACACAACAGCTCTGACTCCAACTGTGATGACAGAAAGAGACCCACCT 2760

Qy     4763 TAAGAAGGAACTGCACATCAGGGGCCTGTGATGTGTGTTGGCACACAGGCCCTTGGAAGC 4822
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2761 TAAGAAGGAACTGCACATCAGGGGCCTGTGATGTGTGTTGGCACACAGGCCCTTGGAAGC 2820

Qy     4823 CCTGTACAGCAGCCTGTGGCAGGGGTTTCCAGTCTGGGAAAGTCGACTGTATCCACACAA 4882
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2821 CCTGTACAGCAGCCTGTGGCAGGGGTTTCCAGTCTGGGAAAGTCGACTGTATCCACACAA 2880

Qy     4883 GGAGTTGCAAACCTGTGGCCAAGAGACACTGTGTACAGAAAAACAAACCAATTTCCTGGC 4942
            ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db     2881 GGAGTTGCAAACCTGTGGCCAAGAGACACTGTGTACAGAAAAAGAAACCAATTTCCTGGC 2940

Qy     4943 GGCACTGTCTTGGGCCCTCCTGTGATAGAGACTGCACAGACACAACTCACTACTGTATGT 5002
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2941 GGCACTGTCTTGGGCCCTCCTGTGATAGAGACTGCACAGACACAACTCACTACTGTATGT 3000

Qy     5003 TTGTAAAACATCTTAATTTGTGTTCTCTAGACCGCTACAAACAAAGGTGCTGCCAGTCAT 5062
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3001 TTGTAAAACATCTTAATTTGTGTTCTCTAGACCGCTACAAACAAAGGTGCTGCCAGTCAT 3060

Qy     5063 GTCAAGAGGGATAA 5076
            ||||||||||||||
Db     3061 GTCAAGAGGGATAA 3074
```

```
RESULT 8
AX235379
LOCUS       AX235379                5107 bp    DNA     linear   PAT 11 SEP 2001
DEFINITION  Sequence 26 from Patent WO0162928.
ACCESSION   AX235379
VERSION     AX235379.1  GI:15593910
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 5107)
  AUTHORS   Vernet,C.A., Fernandes,E., Shimkets,R.A., Macdougall,J. and
            Spaderna,S.K.
  TITLE     Polypeptides and nucleic acids encoding same
  JOURNAL   Patent: WO 0162928-A 26 30-AUG-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1..5107
                     /organism "Homo sapiens"
                     /db_xref "taxon:9606"
BASE COUNT     1444 a   1062 c   1235 g   1366 t
ORIGIN

  Query Match            60.5%;  Score 3072.4;  DB 6;  Length 5107;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 3073;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy     2003 TAGCAGTGTGCTTACATATCCAGACCCAGCAGACAGTCAATGACAGCTTGTGTGATATGG 2062
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1 TAGCAGTGTGCTTACATATCCAGACCCAGCAGACAGTCAATGACAGCTTGTGTGATATGG 60

Qy     2063 TCCACCGTCCTCCAGCCATGAGCCAGGCCTGTAACACAGAGCCCTGTCCCCCCAGGTGGC 2122
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61 TCCACCGTCCTCCAGCCATGAGCCAGGCCTGTAACACAGAGCCCTGTCCCCCCAGGTGGC 120

Qy     2123 ATGTGGGCTCTTGGGGGCCCTGCTCAGCTACCTGTGGAGTTGGAATTCAGACCCGAGATG 2182
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      121 ATGTGGGCTCTTGGGGGCCCTGCTCAGCTACCTGTGGAGTTGGAATTCAGACCCGAGATG 180

Qy     2183 TCTACTGCCTCCACCCAGGGCAGACCCCTCCCCCTCCTGAGGAGTGCCCAGATGAAAAGC 2242
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      181 TCTACTGCCTCCACCCAGGGCAGACCCCTCCCCCTCCTGAGGAGTGCCCAGATGAAAAGC 240

Qy     2243 CCCATGCTTTACAAGCATGCAATCAGTTTGACTGCCCTCCTGGCTGGCACATTGAAGAAT 2302
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241 CCCATGCTTTACAAGCATGCAATCAGTTTGACTGCCCTCCTGGCTGGCACATTGAAGAAT 300

Qy     2303 GGCAGCAGTGTTCCAGGACTTCTGGCGGGGGAACTCAGAACAGAAGAGTCACCTGTCGGC 2362
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      301 GGCAGCAGTGTTCCAGGACTTCTGGCGGGGGAACTCAGAACAGAAGAGTCACCTGTCGGC 360

Qy     2363 AGCTGCTAACGGATGGCAGCTTTTTGAATCTCTCAGATGAATTGTGCCAACCACCCAAGG 2422
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      361 AGCTGCTAACGGATGGCAGCTTTTTGAATCTCTCAGATGAATTGTGCCAACCACCCAAGG 420

Qy     2423 CATCGTCTCACAAGTCCTGTGCCAGGACAGACTGTCCTCCACATTTAGCTGTGGGAGACT 2482
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      421 CATCGTCTCACAAGTCCTGTGCCAGGACAGACTGTCCTCCACATTTAGCTGTGGGAGACT 480

Qy     2483 GGTCGAAGTGTTCTGTCAGTTGTGGTGTTGGAATCCAGAGAAGAAAGCAGGTGTGTCAAA 2542
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      481 GGTCGAAGTGTTCTGTCAGTTGTGGTGTTGGAATCCAGAGAAGAAAGCAGGTGTGTCAAA 540

Qy     2543 GGCTGGCAGCCAAAGGTCGGCGCATCCCCCTCAGTGAGATGATGTGCAGGGATCTACCAG 2602
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      541 GGCTGGCAGCCAAAGGTCGGCGCATCCCCCTCAGTGAGATGATGTGCAGGGATCTACCAG 600

Qy     2603 GCTTCCCTCTTGTAAGATCTTGCCAGATGCCTGAGTGCAGTAAAATCAAATCAGAGATGA 2662
            || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      601 GGCTCCCTCTTGTAAGATCTTGCCAGATGCCTGAGTGCAGTAAAATCAAATCAGAGATGA 660
```

```
Qy       2663 AGACAAAACTTGGTGAGCAGGGTCCGCAGATCCTCAGTGTCCAGAGAGTCTACATTCAGA 2722
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AGACAAAACTTGGTGAGCAGGGTCCGCAGATCCTCAGTGTCCAGAGAGTCTACATTCAGA 720

Qy       2723 CAAGGGAAGAGAAGGGTATTAACCTGACCATTGGTAGCAGAGCCTATTTGCTGCCCAACA 2782
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CAAGGGAAGAGAAGGGTATTAACCTGACCATTGGTAGCAGAGCCTATTTGCTGCCCAACA 780

Qy       2783 CATCCGTGATTATTAAGTGCCCCGTGCGACGATTCCAGAAATCTCTGATCCAGTGGGAGA 2842
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CATCCGTGATTATTAAGTGCCCCGTGCGACGATTCCAGAAATCTCTGATCCAGTGGGAGA 840

Qy       2843 AGGATGGCCGTTTGCTGCAGAACTCCAAACGGCTTGGCATCACCAAGTCAGGCTCACTAA 2902
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AGGATGGCCGTTTGCTGCAGAACTCCAAACGGCTTGGCATCACCAAGTCAGGCTCACTAA 900

Qy       2903 AAATCCAGGGTCTTGCTGCCCCCCGACATCGGGGTGTACGGTGCATTGCAGGCCCTGCAC 2962
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AAATCCAGGGTCTTGCTGCCCCCCGACATCGGGGTGTACGGTGCATTGCAGGCCCTGCAC 960

Qy       2963 AGGAAACAGTTGTGCTCAAGCTCATTGGTACTGACAACCGGCTCATCGCACGCCCAGCCC 3022
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 AGGAAACAGTTGTGCTCAAGCTCATTGGTACTGACAACCGGCTCATCGCACGCCCAGCCC 1020

Qy       3023 TCAGGGAGCCTATGAGGGAATATCCTGGGATGGATCACAGCGAAGCCAATAGTTTGGGAT 3082
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TCAGGGAGCCTATGAGGGAATATCCTGGGATGGATCACAGCGAAGCCAATAGTTTGGGAT 1080

Qy       3083 TCACATGGCACAAAATGAGGCAAATGTGGAATAACAAAAATGACCTTTATCTGGATGATG 3142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TCACATGGCACAAAATGAGGCAAATGTGGAATAACAAAAATGACCTTTATCTGGATGATG 1140

Qy       3143 ACCACATTAGTAACCAGCCTTTCTTGAGAGCTCTGTTAGGCCACTGCAGCAATTCTGCAG 3202
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 ACCACATTAGTAACCAGCCTTTCTTGAGAGCTCTGTTAGGCCACTGCAGCAATTCTGCAG 1200

Qy       3203 GAAGCACCAACTCCTGGGAGTTGAAGAATAAGCAGTTTGAAGCAGCAGTTAAACAAGGAT 3262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GAAGCACCAACTCCTGGGAGTTGAAGAATAAGCAGTTTGAAGCAGCAGTTAAACAAGGAT 1260

Qy       3263 CATATAGCATGGATACAGCCCAGTTTGATGAGCTGATAAGAAACATGAGTCAGCTCATGG 3322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CATATAGCATGGATACAGCCCAGTTTGATGAGCTGATAAGAAACATGAGTCAGCTCATGG 1320

Qy       3323 AAACCGGAGAGGTCAGGGATGATCTTGCGTCCCAGCTGATATATCAGCTGGTGGCCGAAT 3382
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 AAACCGGAGAGGTCAGGGATGATCTTGCGTCCCAGCTGATATATCAGCTGGTGGCCGAAT 1380

Qy       3383 TAGCCAAGGCACAGCCAACACACATGCAGTGGCGGGGCATCCAGGAAGAGACACCTCCTC 3442
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 TAGCCAAGGCACAGCCAACACACATGCAGTGGCGGGGCATCCAGGAAGAGACACCTCCTC 1440

Qy       3443 CTGCTCAGCTCAGAGGGGAAACAGGGAGTGTGTCCCAAAGCTGGCATGCAAAAAACTCAC 3502
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 CTGCTCAGCTCAGAGGGGAAACAGGGAGTGTGTCCCAAAGCTGGCATGCAAAAAACTCAC 1500

Qy       3503 GCAAGCTGACATTCAAGCCGAAAGGACCTGGTTCTCATGAGGCAAAGCCAACCTCCTCAA 3562
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 GCAAGCTGACATTCAAGCCGAAAGGACCTGGTTCTCATGAGGCAAAGCCAACCTCCTCAA 1560

Qy       3563 TTTCATTTAATAAAACAATAAATTCCAGGATTGGAAATACAGTATACATTACAAAAAGGA 3622
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 TTTCATTTAATAAAACAATAAATTCCAGGATTGGAAATACAGTATACATTACAAAAAGGA 1620

Qy       3623 CAGAGGTCATCAATATACTGTGGGACCTTATTACCCCCAGTGAGGCCACATATACATGGA 3682
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 CAGAGGTCATCAATATACTGTGGGACCTTATTACCCCCAGTGAGGCCACATATACATGGA 1680

Qy       3683 CCAAGGATGGAACCTTGTTACAGCCCTCAGTAAAAATAATTTTGGATGGAACTGGGAAGA 3742
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 CCAAGGATGGAACCTTGTTACAGCCCTCAGTAAAAATAATTTTGGATGGAACTGGGAAGA 1740

Qy       3743 TACAGATACAGAATCCTACAAGGAAAGAACAAGGCATATATGAATGTTCTGTAGCTAATC 3802
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 TACAGATACAGAATCCTACAAGGAAAGAACAAGGCATATATGAATGTTCTGTAGCTAATC 1800

Qy       3803 ATCTTGGTTCAGATGTGGAAAGTTCTTCTGTGCTGTATGCAGAGCACCTGGCATCTTGGT 3862
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 ATCTTGGTTCAGATGTGGAAAGTTCTTCTGTGCTGTATGCAGAGCACCTGGCATCTTGGT 1860

Qy       3863 CTGTTGAAAGAAATATCACCAAACCAGAGCACAACCATCTGTCTGTTGTGGTTGGAGGCA 3922
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 CTGTTGAAAGAAATATCACCAAACCAGAGCACAACCATCTGTCTGTTGTGGTTGGAGGCA 1920

Qy       3923 TCGTGGAGGCAGCCCTTGGAGCAAACGTGACAATCCGATGTCCTGTAAAAGGTGTCCCTC 3982
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 TCGTGGAGGCAGCCCTTGGAGCAAACGTGACAATCCGATGTCCTGTAAAAGGTGTCCCTC 1980

Qy       3983 AGCCTAATATAACTTGGTTGAAGAGAGGAGGATCTCTGAGTGGCAATGTTTCCTTGCTTT 4042
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 AGCCTAATATAACTTGGTTGAAGAGAGGAGGATCTCTGAGTGGCAATGTTTCCTTGCTTT 2040

Qy       4043 TCAATGGATCCCTGTTGTTGCAGAATGTTTCCCTTGAAAATGAAGGAACCTACGTCTGCA 4102
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 TCAATGGATCCCTGTTGTTGCAGAATGTTTCCCTTGAAAATGAAGGAACCTACGTCTGCA 2100

Qy       4103 TAGCCACCAATGCTCTTGGAAAGGCAGTGGCAACATCTGTACTTCACTTGCTGGAACGAA 4162
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 TAGCCACCAATGCTCTTGGAAAGGCAGTGGCAACATCTGTACTTCACTTGCTGGAACGAA 2160

Qy       4163 GATGGCCAGAGAGTAGAATCGTATTTCTGCAAGGACATAAAAAGTACATTCTCCAGGCAA 4222
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 GATGGCCAGAGAGTAGAATCGTATTTCTGCAAGGACATAAAAAGTACATTCTCCAGGCAA 2220

Qy       4223 CCAACACTAGAACCAACAGCAATGACCCAACAGGAGAACCCCGGCTCTAAGAGCCTTTTT 4282
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 CCAACACTAGAACCAACAGCAATGACCCAACAGGAGAACCCCGGCTCTAAGAGCCTTTTT 2280

Qy       4283 GGGAGCCTGGTAACTGGTCACATTGTTCTGCCACCTGTGGTCATTTGGGAGCCCGCATTC 4342
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 GGGAGCCTGGTAACTGGTCACATTGTTCTGCCACCTGTGGTCATTTGGGAGCCCGCATTC 2340

Qy       4343 AGAGACCCCAGTGTGTGATGGCCAATGGGAGGAAGTGAGTGAGGCCCCTGTGTGATCACC 4402
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 AGAGACCCCAGTGTGTGATGGCCAATGGGAGGAAGTGAGTGAGGCCCCTGTGTGATCACC 2400

Qy       4403 TCCAGAAGCCACTGGCTGGGTTTGAGCCCTGTAACATCCGGGACCGGCCAGCGAGGTGGT 4462
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 TCCAGAAGCCACTGGCTGGGTTTGAGCCCTGTAACATCCGGGACCGGCCAGCGAGGTGGT 2460

Qy       4463 TCACAAGTGTGTGGTCACAGTGCTCTGTGTCTTGCGGTGAAGGATACCACAGTCGGCAGG 4522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 TCACAAGTGTGTGGTCACAGTGCTCTGTGTCTTGCGGTGAAGGATACCACAGTCGGCAGG 2520

Qy       4523 TGACGTGCAAGCGGACAAAAGCCAATGGAACTGTGCAGGTGGTGTCTCCAACACCATGTG 4582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 TGACGTGCAAGCGGACAAAAGCCAATGGAACTGTGCAGGTGGTGTCTCCAACACCATGTG 2580

Qy       4583 CCCCTAAAGACCGGCCTCTGGGAAGAAAACCATGTTTTGGTCATCCATGTGTTCAGTGGG 4642
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 CCCCTAAAGACCGGCCTCTGGGAAGAAAACCATGTTTTGGTCATCCATGTGTTCAGTGGG 2640

Qy       4643 AACCAGGGAACCGGTGTCCTGGACGTTGCATGGGCCGTGCTGTGAGGATGCAGCAGCGTC 4702
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 AACCAGGGAACCGGTGTCCTGGACGTTGCATGGGCCGTGCTGTGAGGATGCAGCAGCGTC 2700

Qy       4703 ACACAGCTTGTCAACACAACAGTTCTGACTCCAATGTGATGACAGAAAGAGACCCCACCT 4762
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 ACACAGCTTGTCAACACAACAGTTCTGACTCCAATGTGATGACAGAAAGAGACCCCACCT 2760

Qy       4763 TAAGAAGGAACTGCACATCAGGGCCCCTGTGATGTGGTTGGCACACAGGCCCTTGGAAGC 4822
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 TAAGAAGGAACTGCACATCAGGGCCCCTGTGATGTGGTTGGCACACAGGCCCTTGGAAGC 2820

Qy       4823 CCTGTACAGAGCCTGTGGCAGCAGTTTCCAGCTTTTGGAAAGCGACTGTATCCCACACAA 4882
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 CCTGTACAGCAGCCTGTGGCAGGGGTTTCCAGTCTCGGAAAGTCGACTGTATCCACACAA 2880

Qy       4883 GGAGTTGCAAACCTGTGGCCAAGAGACACTGTGTACAGAAAAAGAAACCAATTTCCTGGC 4942
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 GGAGTTGCAAACCTGTGGCCAAGAGACACTGTGTACAGAAAAAGAAACCAATTTCCTGGC 2940

Qy       4943 GGCACTGTCTTGGGCCCTCCTGTGATAGAGACTGCACAGACACAACTCACTACTGTATGT 5002
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 GGCACTGTCTTGGGCCCTCCTGTGATAGAGACTGCACAGACACAACTCACTACTGTATGT 3000

Qy       5003 TTGTAAAACATCTTAATTTGTGTTCTCTAGACCGCTACAAACAAAGGTGCTGCCAGTCAT 5062
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 TTGTAAAACATCTTAATTTGTGTTCTCTAGACCGCTACAAACAAAGGTGCTGCCAGTCAT 3060

Qy       5063 GTCAAGAGGGATAA 5076
              ||||||||||||||
Db       3061 GTCAAGAGGGATAA 3074


RESULT 9
AX235413
LOCUS       AX235413                5107 bp    DNA     linear   PAT 11-SEP-2001
DEFINITION  Sequence 60 from Patent WO0162928.
ACCESSION   AX235413
VERSION     AX235413.1  GI:15593926
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 5107)
  AUTHORS   Vernet,C.A., Fernandes,E., Shimkets,R.A., Macdougall,J. and
            Spaderna,S.K.
  TITLE     Polypeptides and nucleic acids encoding same
  JOURNAL   Patent: WO 0162928-A 60 30-AUG-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1..5107
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT     1444 a   1061 c   1235 g   1367 t
ORIGIN

  Query Match             60.5%;  Score 3072.4;  DB 6;  Length 5107;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 3073;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2003 TAGCAGTGTGCTTACATATCCAGACCCAGCAGACAGTCAATGACAGCTTGTGTGATATGG 2062
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TAGCAGTGTGCTTACATATCCAGACCCAGCAGACAGTCAATGACAGCTTGTGTGATATGG 60

Qy       2063 TCCACCGTCCTCCAGCCATGAGCCAGGCCTGTAACACAGAGCCCTGTCCCCCCAGGTGGC 2122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCCACCGTCCTCCAGCCATGAGCCAGGCCTGTAACACAGAGCCCTGTCCCCCCAGGTGGC 120

Qy       2123 ATGTGGGCTCTTGGGGGCCCTGCTCAGCTACCTGTGCAGTTGGAATTCAGACCCGAGATG 2182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ATGTGGGCTCTTGGGGGCCCTGCTCAGCTACCTGTGGAGTTGGAATTCAGACCCGAGATG 180

Qy       2183 TGTACTGCCTGCACCCAGGGGAGACCCCTGCCCCTCCTGAGGAGTGCCGAGATGAAAAGC 2242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGTACTGCCTGCACCCAGGGGAGACCCCTGCCCCTCCTGAGGAGTGCCGAGATGAAAAGC 240

Qy       2243 CCCATGCTTTACAAGCATGCAATCAGTTTGACTGCCCTCCTGCCTGGCACATTGAAGAAT 2302
              |||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db        241 CCCATGCTTTACAAGCATGCAATCAGTTTGACTGCCCTCCTGGCTGGCACATTGAAGAAT 300

Qy       2303 GGCAGCAGTGTTCCAGGACTTGTGGCGGGGAACTCAGAACAGAAGAGTCACCTGTCGGC 2362
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGCAGCAGTGTTCCAGGACTTGTGGCGGGGAACTCAGAACAGAAGAGTCACCTGTCGGC 360

Qy       2363 AGCTGCTAACGGATGGCAGCTTTTTGAATCTCTCAGATGAATTGTGCCAAGGACCCAAGG 2422
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AGCTGCTAACGGATGGCAGCTTTTTGAATCTCTCAGATGAATTGTGCCAAGGACCCAAGG 420

Qy       2423 CATCGTCTCACAAGTCCTGTGCCAGGACAGACTGTCCTCCACATTTAGCTGTGGGAGACT 2482
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CATCGTCTCACAAGTCCTGTGCCAGGACAGACTGTCCTCCACATTTAGCTGTGGGAGACT 480

Qy       2483 GGTCGAAGTGTTCTGTCAGTTGTGGTGTTGGAATCCAGAGAAGAAAGCAGGTGTGTCAAA 2542
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GGTCGAAGTGTTCTGTCAGTTGTGGTGTTGGAATCCAGAGAAGAAAGCAGGTGTGTCAAA 540

Qy       2543 GGCTGGCAGCCAAAGGTCGGCCCATCCCCCTCAGTGAGATCATGTGCAGGGATCTACCAG 2602
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GGCTGGCAGCCAAAGGTCGGCCCATCCCCCTCAGTGAGATCATGTGCAGGGATCTACCAG 600

Qy       2603 GGTTCCCTCTTGTAAGATCTTGCCAGATGCCTGAGTGCAGTAAAATCAAATCAGAGATGA 2662
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GGTTCCCTCTTGTAAGATCTTGCCAGATGCCTGAGTGCAGTAAAATCAAATCAGAGATGA 660

Qy       2663 AGACAAAACTTGGTGAGCAGGGTCCGCAGATCCTCAGTGTCCAGAGAGTCTACATTCAGA 2722
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AGACAAAACTTGGTGAGCAGGGTCCGCAGATCCTCAGTGTCCAGAGAGTCTACATTCAGA 720

Qy       2723 CAAGGGAAGAGAAGCGTATTAACCTGACCATTGGTAGCAGAGCCTATTTGCTGCCCAACA 2782
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CAAGGGAAGAGAAGCGTATTAACCTGACCATTGGTAGCAGAGCCTATTTGCTGCCCAACA 780

Qy       2783 CATCCGTGATTATTAAGTGCCCCGTGCGACGATTCCAGAAATCTCTGATCCAGTGGGAGA 2842
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CATCCGTGATTATTAAGTGCCCCGTGCGACGATTCCAGAAATCTCTGATCCAGTGGGAGA 840

Qy       2843 AGGATGGCCGTTGCCTGCAGAACTCCAAACGGCTTGGCATCACCAAGTCAGGCTCACTAA 2902
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AGGATGGCCGTTGCCTGCAGAACTCCAAACGGCTTGGCATCACCAAGTCAGGCTCACTAA 900

Qy       2903 AAATCCACGGTCTTGCTGCCCCCGACATCGGCGTGTACCGGTGCATTGCAGGCTCTGCAC 2962
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AAATCCACGGTCTTGCTGCCCCCGACATCGGCGTGTACCGGTGCATTGCAGGCTCTGCAC 960

Qy       2963 AGGAAACAGTTGTGCTCAAGCTCATTGGTACTGACAACCGGCTCATCGCACGCCCAGCCC 3022
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 AGGAAACAGTTGTGCTCAAGCTCATTGGTACTGACAACCGGCTCATCGCACGCCCAGCCC 1020

Qy       3023 TCAGGGAGCCTATGAGGGAATATCCTGGGATGGACCACAGCGAAGCCAATAGTTTGGGAG 3082
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TCAGGGAGCCTATGAGGGAATATCCTGGGATGGACCACAGCGAAGCCAATAGTTTGGGAG 1080

Qy       3083 TCACATGGCACAAAATGAGGCAAATGTGGAATAACAAAAATGACCTTTATCTGGATGATG 3142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TCACATGGCACAAAATGAGGCAAATGTGGAATAACAAAAATGACCTTTATCTGGATGATG 1140

Qy       3143 ACCACATTAGTAACCAGCCTTTCTTGAGAGCTCTGTTAGGCCACTGCAGCAATTCTGCAG 3202
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 ACCACATTAGTAACCAGCCTTTCTTGAGAGCTCTGTTAGGCCACTGCAGCAATTCTGCAG 1200

Qy       3203 GAAGCACCAACTCCTGGGAGTTGAAGAATAAGCACTTTGAAGCAGCAGTTAAACAAGGAG 3262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GAAGCACCAACTCCTGGGAGTTGAAGAATAAGCACTTTGAAGCAGCAGTTAAACAAGGAG 1260

Qy       3263 CATATAGCATGGATACAGCCCAGTTTGATGAGCTGATAAGAAACATGAGTCAGCTCATGG 3322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CATATAGCATGGATACAGCCCAGTTTGATGAGCTGATAAGAAACATGAGTCAGCTCATGG 1320

Qy       3323 AAACCGGAGAGGTCAGCGATGATCTTGCGTCCCAGCTGATATATCAGCTGGTGGCCGAAT 3382
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 AAACCGGAGAGGTCAGCGATGATCTTGCGTCCCAGCTGATATATCAGCTGGTGGCCGAAT 1380

Qy       3383 TAGCCAAGGCACAGCCAACACACATGCAGTGGCGGGGCATCCAGGAAGAGACACCTCCTG 3442
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 TAGCCAAGGCACAGCCAACACACATGCAGTGGCGGGGCATCCAGGAAGAGACACCTCCTG 1440

Qy       3443 CTGCTCAGCTCAGAGGGGAAACAGGGAGTGTGTCCCAAAGCTCGGCATGCAAAAACTCAG 3502
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 CTGCTCAGCTCAGAGGGGAAACAGGGAGTGTGTCCCAAAGCTCGGCATGCAAAAACTCAG 1500
```

```
Qy       3503 GCAAGCTGACATTCAAGCTGAAAGGACCTGTTCTCATGAGGCAAAGCCAACCTCCCTCAA 3562
Db       1501 GCAAGCTGACATTCAAGCTGAAAGGACCTGTTCTCATGAGGCAAAGCCAACCTCCCTCAA 1560

Qy       3563 TTTCATTTAATAAAACAATAAATTCCAGGATTGGAAATACAGTATACATTACAAAAAGGA 3622
Db       1561 TTTCATTTAATAAAACAATAAATTCCAGGATTGGAAATACAGTATACATTACAAAAAGGA 1620

Qy       3623 CAGAGGTCATCAATATACTGGTGACCTTATTACCCCCAGTGAGGCCACATATACATGGA 3682
Db       1621 CAGAGGTCATCAATATACTGGTGACCTTATTACCCCCAGTGAGGCCACATATACATGGA 1680

Qy       3683 CCAAGGATGGAACCTTGTTACAGCCCTCAGTAAAAATAATTTTGGATGGAACTGGGAAGA 3742
Db       1681 CCAAGGATGGAACCTTGTTACAGCCCTCAGTAAAAATAATTTTGGATGGAACTGGGAAGA 1740

Qy       3743 TACAGATACAGAATCCTACAAGGAAAGAACAAGGCATATATGAATGTTCTGTAGCTAATC 3802
Db       1741 TACAGATACAGAATCCTACAAGGAAAGAACAAGGCATATATGAATGTTCTGTAGCTAATC 1800

Qy       3803 ATCTTGGTTCAGATGTGGAAAGTTCTTCTGTGCTGTATGCAGAGGCACCTGTCATCTTGT 3862
Db       1801 ATCTTGGTTCAGATGTGGAAAGTTCTTCTGTGCTGTATGCAGAGGCACCTGTCATCTTGT 1860

Qy       3863 CTGTTGAAAGAAATATCACCAAACCAGAGCACAACCATCTGTCTGTTGTGGCTGGAGGCA 3922
Db       1861 CTGTTGAAAGAAATATCACCAAACCAGAGCACAACCATCTGTCTGTTGTGGCTGGAGGCA 1920

Qy       3923 TGGTGGAGGCAGCCTTGGAGCAAACGTGACAATCCGATGTCCTGTAAAAGGTGTCCCTC 3982
Db       1921 TGGTGGAGGCAGCCTTGGAGCAAACGTGACAATCCGATGTCCTGTAAAAGGTGTCCCTC 1980

Qy       3983 AGCCTAATATAACTTGGTGGAAGAGAGGAGGATCTCTGAGTGGCAATGTTTCCTTGCTTT 4042
Db       1981 AGCCTAATATAACTTGGTGGAAGAGAGGAGGATCTCTGAGTGGCAATGTTTCCTTGCTTT 2040

Qy       4043 TCAATGGATCCCTGCTGTTGCAGAATGTTTCCCTTGAAAATGAAGGAACCTACGTCTGCA 4102
Db       2041 TCAATGGATCCCTGCTGTTGCAGAATGTTTCCCTTGAAAATGAAGGAACCTACGTCTGCA 2100

Qy       4103 TAGCCACCAATGCTCTTGGAAAGGCAGTGGCAACATCTGTATTCCACTTGCTGGAACGAA 4162
Db       2101 TAGCCACCAATGCTCTTGGAAAGGCAGTGGCAACATCTGTATTCCACTTGCTGGAACGAA 2160

Qy       4163 GATGGCCAGAGAGTAGAATCGTATTTCTGCAAGGACATAAAAAGTACATTCTCCAGGCAA 4222
Db       2161 GATGGCCAGAGAGTAGAATCGTATTTCTGCAAGGACATAAAAAGTACATTCTCCAGGCAA 2220

Qy       4223 CCAACACTAGAACCAACAGCAATGACCCAACAGGAGAACCCCGCCTCAAGAGCCTTTTT 4282
Db       2221 CCAACACTAGAACCAACAGCAATGACCCAACAGGAGAACCCCGCCTCAAGAGCCTTTTT 2280

Qy       4283 GGGAGCCTGGTAACTGGTCACATTGTTCTGCCACCTGTGGTCATTGGGAGCCCGCATTC 4342
Db       2281 GGGAGCCTGGTAACTGGTCACATTGTTCTGCCACCTGTGGTCATTGGGAGCCCGCATTC 2340

Qy       4343 AGAGACCCAGTGTGTGATGGCCAATGGGCAGGAAGTGAGTGAGGCCCTGTGTGATCACC 4402
Db       2341 AGAGACCCAGTGTGTGATGGCCAATGGGCAGGAAGTGAGTGAGGCCCTGTGTGATCACC 2400

Qy       4403 TCCAGAAGCCACTGGCTGGGTTTGAGCCTCCTAACATCCGGGACTGGCCAGCGAGGTGGT 4462
Db       2401 TCCAGAA[illegible] [illegible]

Qy       4463 TCACAAGTGTGTGGTCACAGGCTCTGTGCTTGCGGTGAAGGATACCACAGTCGGCAGG 4522
Db       2461 TCACAAGTGTGTGGTCACAGGCTCTGTGCTTGCGGTGAAGGATACCACAGTCGGCAGG 2520

Qy       4523 TGACGTGCAAGCGGACAAAAGCCAATGGAACTGTGCAGGTGGTGTCTCCAAGAGCATGTG 4582
Db       2521 TGACGTGCAAGCGGACAAAAGCCAATGGAACTGTGCAGGTGGTGTCTCCAAGAGCATGTG 2580

Qy       4583 CCCCTAAAGACCGGCCTCTGGAAGAAAACCATGTTTTGGTCATCCATGTTTCAGTGGG 4642
Db       2581 CCCCTAAAGACCGGCCTCTGGAAGAAAACCATGTTTTGGTCATCCATGTTTCAGTGGG 2640

Qy       4643 AACCAGGGAACCGGTGTCCTGGACGTTGCATGGGCGGTGCTGTGAGGATGCAGCAGGCTC 4702
Db       2641 AACCAGGGAACCGGTGTCCTGGACGTTGCATGGGCGGTGCTGTGAGGATGCAGCAGGCTC 2700

Qy       4703 ACACAGCTTGTCAACACAACAGCTCTGACTCCAACTGTGATGACAGAAAGAGACCCACCT 4762
Db       2701 ACACAGCTTGTCAACACAACAGCTCTGACTCCAACTGTGATGACAGAAAGAGACCCACCT 2760

Qy       4763 TAAGAAGGAACTGCACATCAGGGGCCTGTGATGTGTGTGGCACACAGGCCCTTGGAAGC 4822
Db       2761 TAAGAAGGAACTGCACATCAGGGGCCTGTGATGTGTGTGGCACACAGGCCCTTGGAAGC 2820

Qy       4823 CCTGTACAGCAGCCTGTGGCAGGGTTTCCAGTCTCGGAAAGTGGACTGGTATCCACACAA 4882
Db       2821 CCTGTACAGCAGCCTGTGGCAGGGTTTCCAGTCTCGGAAAGTGGACTGGTATCCACACAA 2880

Qy       4883 GGAGTTGCAAACTGTGGCCAAGAGACACTGTGTACAGAAAAAGAAACCAATTTCCTGGC 4942
Db       2881 GGAGTTGCAAACTGTGGCCAAGAGACACTGTGTACAGAAAAAGAAACCAATTTCCTGGC 2940

Qy       4943 GGCACTGTCTTGGACCCTCCTGGATAGAGACTGCACAGACACAACTCACTACTGTATGT 5002
Db       2941 GGCACTGTCTTGGACCCTCCTGGATAGAGACTGCACAGACACAACTCACTACTGTATGT 3000

Qy       5003 TTGTAAAACATCTTAATTTGGGTTCTCTAGACCCTACAAACAAAGGGCTGGCCAGTCAT 5062
Db       3001 TTGTAAAACATCTTAATTTGGGTTCTCTAGACCCTACAAACAAAGGGCTGGCCAGTCAT 3060

Qy       5063 GTCAAGAGGGATAA 5076
Db       3061 GTCAAGAGGGATAA 3074


RESULT 10
AB033059
LOCUS       AB033059                5107 bp    mRNA    linear   PRI 11-NOV-1999
DEFINITION  Homo sapiens mRNA for KIAA1233 protein, partial cds.
ACCESSION   AB033059
VERSION     AB033059.1  GI:6330728
KEYWORDS    .
SOURCE      Homo sapiens brain cDNA to mRNA, clone_lib:pBluescriptII SK plus
            clone:fh08611.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (sites)
  AUTHORS   Nagase,T., Ishikawa,K., Kikuno,R., Hirosawa,M., Nomura,N. and
            Ohara,O.
  TITLE     Prediction of the coding sequences of unidentified human genes. XV.
            The complete sequences of 100 new cDNA clones from brain which code
            for large proteins in vitro
  JOURNAL   DNA Res. 6 (5), 337-345 (1999)
  MEDLINE   20039619
REFERENCE   2  (bases 1 to 5107)
  AUTHORS   Ohara,O., Nagase,T. and Kikuno,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-OCT-1999) Osamu Ohara, Kazusa DNA Research Institute,
            Laboratory of DNA Technology; 1532-3 Yana, Kisarazu, Chiba
            292-0812, Japan (E-mail:cdnainfo@kazusa.or.jp,
            URL:http://www.kazusa.or.jp/huge/, Tel:+81-438-52-3913,
            Fax:+81-438-52-3914)
FEATURES             Location/Qualifiers
     source          1..5107
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="fh08611"
                     /tissue_type="brain"
                     /clone_lib="pBluescriptII SK plus"
     gene            1..5107
```

```
                     /gene "KIAA1233"
     CDS             <1..3074
                     /gene "KIAA1233"
                     /note "Start codon is not identified."
                     /codon_start 3
                     /product="KIAA1233 protein"
                     /protein_id="BAA86547.1"
                     /db_xref="GI:6330729"
                     /translation="AVCLHIQTLQTVNFSLCDMVHEPFAMSQACNTEPCPFEWHVGEW
                     GPCSATCGVGIQTRDVYCLHPGETPAPPEECRDEKPHALQACNQFDCPPGWHIEEWQQ
                     CSRTCGGGTQNRRVTCRQLLTDGSFLNLSDELCQGPKASSHKSCARTDCPPHLAVGDW
                     SKCSVSCGVGIQRRKQVCQRLAAKGRRIPLSEMMCRDLPGLPLVRSCQMPECSKIKSE
                     MKTKLGEQGPQILSVQRVYIQTREEKRINLTIGSRAYLLPNTSVIIKCPVRRFQKSLI
                     QWEKDGRCLQNSKRLGITKSGSLKIHGLAAPDIGVYRCIAGSAQETVVLKLIGTDNRL
                     IARPALREPMREYPGMDHSEANSLGVTWHKMRQMWNNKNDLYLDDDHISNQPFLRALL
                     GHCSNSACSINSWELENKQYLAAVEQGAYSLIAQFDSLLIRNMSQLMEICEVSDDLAS
                     QLIYQLVAELAKAQPTHMQWRCIQEETPPAAQIRCFICSVSQSSHA*NSCKLIEEPFG
                     PVLMRQSQPPSISFNKTINSRIGNTVYITKRTEVINILCDLITPSEATYTWTKDGTLL
                     QPSVKIIIDGTGKIQIQNPTRKEQGIYECSVANHLGSDVESSSVLYAEAPVILSVERN
                     ITKPEHNHLSVVVGGIVEAALGANVTIRCPVKGVPQPNITWLKRGGSLSGNVSLLFNG
                     SLLLQNVSLENEGTYVCIATNALGKAVATSVLHLLERRWPESRIVFLQGHKKYILQAT
                     NTRTNSNDPTGEPPPQEPFWEPGNWSHCSATCGHLGARIQRPQCVMANGQEVSEALCD
                     HLQKPLAGFEPCNIRDCPARWFTSVWSQCSVSCGEGYHSRQVTCKRTKANGTVQVVSP
                     RACAPKDRPLGRKPCFGHPCVQWEPGNRCPGRCMGRAVRMQQRHTACQHNSSDSNCDD
                     RKRPTLRRNCTSGACDVCWHTGPWRPCTAACGRGFQSRKVDCIHTRSCKPVAKRHCVQ
                     KKKPISWRHCLGPSCDRDCTDTTHYCMFVKHLNLCSLDRYKQRCCQSCQEG"
BASE COUNT     1444 a   1062 c   1235 g   1366 t
ORIGIN

  Query Match             60.5%;  Score 3072.4;  DB 9;  Length 5107;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 3073;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2003 TAGCAGTGTGCTTACATATCCAGACCCAGCAGACAGTCAATGACAGCTTGTGTGATATGG 2062
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TAGCAGTGTGCTTACATATCCAGACCCAGCAGACAGTCAATGACAGCTTGTGTGATATGG 60

Qy       2063 TCCACCGTCCTCCAGCCATGAGCCAGGCCTGTAACACAGAGCCCTGTCCCCCCAGGTGGC 2122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCCACCGTCCTCCAGCCATGAGCCAGGCCTGTAACACAGAGCCCTGTCCCCCCAGGTGGC 120

Qy       2123 ATGTGGGCTCTTGGGGGCCCTGCTCAGCTACCTGTGGAGTTGGAATTCAGACCCGAGATG 2182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ATGTGGGCTCTTGGGGGCCCTGCTCAGCTACCTGTGGAGTTGGAATTCAGACCCGAGATG 180

Qy       2183 TGTACTGCCTGCACCCAGGGGAGACCCCTGCCCCTCCTGAGGAGTGCCGAGATGAAAAGC 2242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGTACTGCCTGCACCCAGGGGAGACCCCTGCCCCTCCTGAGGAGTGCCGAGATGAAAAGC 240

Qy       2243 CCCATGCTTTACAAGCATGCAATCAGTTTGACTGCCCTCCTGGCTGGCACATTGAAGAAT 2302
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCCATGCTTTACAAGCATGCAATCAGTTTGACTGCCCTCCTGGCTGGCACATTGAAGAAT 300

Qy       2303 GGCAGCAGTGTTCCAGGACTTGTGGCGGGGGAACTCAGAACAGAAGAGTCACCTGTCGGC 2362
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGCAGCAGTGTTCCAGGACTTGTGGCGGGGGAACTCAGAACAGAAGAGTCACCTGTCGGC 360

Qy       2363 AGCTGCTAACGGATGGCAGCTTTTTGAATCTCTCAGATGAATTGTGCCAAGGACCCAAGG 2422
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AGCTGCTAACGGATGGCAGCTTTTTGAATCTCTCAGATGAATTGTGCCAAGGACCCAAGG 420

Qy       2423 CATCGTCTCACAAGTCCTGTGCCAGGACAGACTGTCCTCCACATTTAGCTGTGGGAGACT 2482
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CATCGTCTCACAAGTCCTGTGCCAGGACAGACTGTCCTCCACATTTAGCTGTGGGAGACT 480

Qy       2483 GGTCGAAGTGTTCTGTCAGTTGTGGTGTTGGAATCCAGAGAAGAAAGCAGGTGTGTCAAA 2542
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GGTCGAAGTGTTCTGTCAGTTGTGGTGTTGGAATCCAGAGAAGAAAGCAGGTGTGTCAAA 540

Qy       2543 GGCTGGCAGCCAAAGGTCGGCGCATCCCCCTCAGTGAGATGATGTGCAGGGATCTACCAG 2602
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GGCTGGCAGCCAAAGGTCGGCGCATCCCCCTCAGTGAGATGATGTGCAGGGATCTACCAG 600

Qy       2603 GGTTCCCTCTTGTAAGATCTTGCCAGATGCCTGAGTGCAGTAAAATCAAATCAGAGATGA 2662
              || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GGCTCCCTCTTGTAAGATCTTGCCAGATGCCTGAGTGCAGTAAAATCAAATCAGAGATGA 660

Qy       2663 AGACAAAACTTGGTGAGCAGGGTCCGCAGATCCTCAGTGTCCAGAGAGTCTACATTCAGA 2722
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AGACAAAACTTGGTGAGCAGGGTCCGCAGATCCTCAGTGTCCAGAGAGTCTACATTCAGA 720

Qy       2723 CAAGGCAAGACAACCCTATTAACCTCACCATTGGTACCAGAGCCTATTTCCTGCCCAACA 2782
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CAAGGCAAGACAACCCTATTAACCTCACCATTGGTACCAGAGCCTATTTCCTGCCCAACA 780

Qy       2783 CATCCGTGATTATTAAGTGCCCCGTGCGACGATTCCAGAAATCTCTGATCCAGTGGGAGA 2842
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CATCCGTGATTATTAAGTGCCCCGTGCGACGATTCCAGAAATCTCTGATCCAGTGGGAGA 840

Qy       2843 AGGATGGGCGTTGCCTGCAGAACTCCAAACGGCTTGGCATCACCAAGTCAGGCTCACTAA 2902
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AGGATGGGCGTTGCCTGCAGAACTCCAAACGGCTTGGCATCACCAAGTCAGGCTCACTAA 900

Qy       2903 AAATCCACGGTCTTGCTGCCCCCGACATCGGCGTGTACCGGTGCATTGCAGGCTCTGCAC 2962
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AAATCCACGGTCTTGCTGCCCCCGACATCGGCGTGTACCGGTGCATTGCAGGCTCTGCAC 960

Qy       2963 AGGAAACAGTTGTGCTCAAGCTCATTGGTACTGACAACCGGCTCATCGCACGGCCAGCCC 3022
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 AGGAAACAGTTGTGCTCAAGCTCATTGGTACTGACAACCGGCTCATCGCACGGCCAGCCC 1020

Qy       3023 TCAGGGAGCCTATGAGGGAATATCCTGGGATGGACCACAGCGAAGCCAATAGTTTGGGAG 3082
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TCAGGGAGCCTATGAGGGAATATCCTGGGATGGACCACAGCGAAGCCAATAGTTTGGGAG 1080

Qy       3083 TCACATGGCACAAAATGAGGCAAATGTGGAATAACAAAAATGACCTTTATCTGGATGATG 3142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TCACATGGCACAAAATGAGGCAAATGTGGAATAACAAAAATGACCTTTATCTGGATGATG 1140

Qy       3143 ACCACATTAGTAACCAGCCTTTCTTGAGAGCTCTGTTAGGCCACTGCAGCAATTCTGCAG 3202
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 ACCACATTAGTAACCAGCCTTTCTTGAGAGCTCTGTTAGGCCACTGCAGCAATTCTGCAG 1200

Qy       3203 GAAGCACCAACTCCTGGGAGTTGAAGAATAAGCAGTTTGAAGCAGCAGTTAAACAAGGAG 3262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GAAGCACCAACTCCTGGGAGTTGAAGAATAAGCAGTTTGAAGCAGCAGTTAAACAAGGAG 1260

Qy       3263 CATATAGCATGGATACAGCCCAGTTTGATGAGCTGATAAGAAACATGAGTCAGCTCATGG 3322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CATATAGCATGGATACAGCCCAGTTTGATGAGCTGATAAGAAACATGAGTCAGCTCATGG 1320

Qy       3323 AAACCGGAGAGGTCAGCGATGATCTTGCGTCCCAGCTGATATATCAGCTGGTGGCCGAAT 3382
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 AAACCGGAGAGGTCAGCGATGATCTTGCGTCCCAGCTGATATATCAGCTGGTGGCCGAAT 1380

Qy       3383 TAGCCAAGGCACAGCCAACACACATGCAGTGGCGGGGCATCCAGGAAGAGACACCTCCTG 3442
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 TAGCCAAGGCACAGCCAACACACATGCAGTGGCGGGGCATCCAGGAAGAGACACCTCCTG 1440

Qy       3443 CTGCTCAGCTCAGAGGGGAAACAGGGAGTGTGTCCCAAAGCTCGCATGCAAAAAACTCAG 3502
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 CTGCTCAGCTCAGAGGGGAAACAGGGAGTGTGTCCCAAAGCTCGCATGCAAAAAACTCAG 1500

Qy       3503 GCAAGCTGACATTCAAGCCGAAAGGACCTGTTCTCATGAGGCAAAGCCAACCTCCCTCAA 3562
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 GCAAGCTGACATTCAAGCCGAAAGGACCTGTTCTCATGAGGCAAAGCCAACCTCCCTCAA 1560

Qy       3563 TTTCATTTAATAAAACAATAAATTCCAGGATTGGAAATACAGTATACATTACAAAAAGGA 3622
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 TTTCATTTAATAAAACAATAAATTCCAGGATTGGAAATACAGTATACATTACAAAAAGGA 1620

Qy       3623 CAGAGGTCATCAATATACTGTGTGACCTTATTACCCCAGTGAGGCCACATATACATGGA 3682
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 CAGAGGTCATCAATATACTGTGTGACCTTATTACCCCAGTGAGGCCACATATACATGGA 1680

Qy       3683 CCAAGGATGGAACCTTGTTACAGCCCTCAGTAAAAATAATTTTGGATGGAACTGGGAAGA 3742
```

```
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 CCAAGGATGGAACCTTGTTACAGGCCTCAGTAAAAATAATTTGGATGGAACTGGGAAGA 1740

Qy       3743 TACAGATACAGAATCCTACAAGGAAAGAACAAGGCATATATGAATGTTCTGTAGCTAATC 3802
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 TACAGATACAGAATCCTACAAGGAAAGAACAAGGCATATATGAATGTTCTGTAGCTAATC 1800

Qy       3803 ATCTTGGTTCAGATGTGGAAATGTCTTCTGTGCTGTATGCAGAGGTACCTGTCATCTTGT 3862
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 ATCTTGGTTCAGATGTGGAAATGTCTTCTGTGCTGTATGCAGAGGTACCTGTCATCTTGT 1860

Qy       3863 CTGTTGAAAGAAATATCACCAAACCAGAGCACAACCATCTGTCTGTTGTGGTTGGAGGCA 3922
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 CTGTTGAAAGAAATATCACCAAACCAGAGCACAACCATCTGTCTGTTGTGGTTGGAGGCA 1920

Qy       3923 TCGTGGAGGCAGCCTTCGGAGAAAGGTGACAATCCGATGTCCTGTAAAAGGTGTCCCTCC 3982
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 TCGTGGAGGCAGCCTTCGGAGAAAGGTGACAATCCGATGTCCTGTAAAAGGTGTCCCTCC 1980

Qy       3983 AGCCTAATATAACTTGGTTGAAGAGAGGAGGATCTCTGAGTGGCAATGTTTCCTTGCTTT 4042
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 AGCCTAATATAACTTGGTTGAAGAGAGGAGGATCTCTGAGTGGCAATGTTTCCTTGCTTT 2040

Qy       4043 TCAATGGATCCCTGGTTGTTCCAGAATGTTTCCCTGAAAATGAAGGAACCTACGTCTGTA 4102
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 TCAATGGATCCCTGGTTGTTCCAGAATGTTTCCCTGAAAATGAAGGAACCTACGTCTGTA 2100

Qy       4103 TAGCCACCAATGCTCTTGGAAAGGCAGTGGCAACATCTGTACTCCACTTGCTGGAACGAA 4162
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 TAGCCACCAATGCTCTTGGAAAGGCAGTGGCAACATCTGTACTCCACTTGCTGGAACGAA 2160

Qy       4163 GATGGCCAGAGAGTAGAATCCTATTTCTGCAAGGACATAAAAAGTACATTCTCCAGGCAA 4222
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 GATGGCCAGAGAGTAGAATCCTATTTCTGCAAGGACATAAAAAGTACATTCTCCAGGCAA 2220

Qy       4223 CCAACACTAGAACCAACAGCAATGACCCAACAGGAGAACCCCGGCCTCAAGAGCCTTTTT 4282
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 CCAACACTAGAACCAACAGCAATGACCCAACAGGAGAACCCCGGCCTCAAGAGCCTTTTT 2280

Qy       4283 GGGAGCCTGGTAACTGGTCACATTGTTCTGCCACCTGTGGTCATTTGGGAGCCCGCATTC 4342
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 GGGAGCCTGGTAACTGGTCACATTGTTCTGCCACCTGTGGTCATTTGGGAGCCCGCATTC 2340

Qy       4343 AGAGACCCCAGTGTGTGATGGCCAATGGGCAGGAAGTGAGTGAGGCCCTGTGTGATCACC 4402
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 AGAGACCCCAGTGTGTGATGGCCAATGGGCAGGAAGTGAGTGAGGCCCTGTGTGATCACC 2400

Qy       4403 TCCAGAAGCCACTGGCTGGGTTTGAGCCCTGTAACATCCGGGACTGGCCAGGGAGGTGGT 4462
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 TCCAGAAGCCACTGGCTGGGTTTGAGCCCTGTAACATCCGGGACTGGCCAGGGAGGTGGT 2460

Qy       4463 TCACAAGTGTGTGGTCACAGTGCTCTGTGTCTTGCGGTGAAGGATACCACAGTCGGCAGG 4522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 TCACAAGTGTGTGGTCACAGTGCTCTGTGTCTTGCGGTGAAGGATACCACAGTCGGCAGG 2520

Qy       4523 TGACGTGCAAGCGGACAAAAGCCAATGGAACTGTGCAGGTGGTGGTTCCAAGAGCATGTT 4582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 TGACGTGCAAGCGGACAAAAGCCAATGGAACTGTGCAGGTGGTGGTTCCAAGAGCATGTT 2580

Qy       4583 CCCCTAAAGACCGGCCTCTGGGAAGAAAACCATGTTTTGGTCATCCATGTGTTCAGTGGG 4642
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 CCCCTAAAGACCGGCCTCTGGGAAGAAAACCATGTTTTGGTCATCCATGTGTTCAGTGGG 2640

Qy       4643 AACCAGGGAACCGGTGTCCTGGACGTTGCATGGGCCGTGCTGTGAGGATGCAGCAGCGCC 4702
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 AACCAGGGAACCGGTGTCCTGGACGTTGCATGGGCCGTGCTGTGAGGATGCAGCAGCGCC 2700

Qy       4703 ACACATTTGTTCAACACAACACCTCTGACTCCAACTGTGATGACAGAAAGAGACCCACCT 4762
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 ACACATTTGTTCAACACAACACCTCTGACTCCAACTGTGATGACAGAAAGAGACCCACCT 2760

Qy       4763 TAAGAAGGAACTGCACATCAGGGGCTGTGATGTTGTTCCCATACAGGCCCCTTTGGAAGC 4822
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 TAAGAAGGAACTGCACATCAGGGGCTGTGATGTTGTTCCCATACAGGCCCCTTTGGAAGC 2820

Qy       4823 CCTGTACAGCAGCCTGTGGCAGGGGTTTCCAGTCTGGAAAGCCGACTGGTATCCACACAA 4882
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 CCTGTACAGCAGCCTGTGGCAGGGGTTTCCAGTCTGGAAAGCCGACTGGTATCCACACAA 2880

Qy       4883 GGAGTTGCAAACCTGTGGCAAGGAGACACTGTGTACAGAAAAAGAAACCAATTTCCTGGC 4942
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 GGAGTTGCAAACCTGTGGCAAGGAGACACTGTGTACAGAAAAAGAAACCAATTTCCTGGC 2940

Qy       4943 GGCACTGTCTTGGGCCCTCCTGTGATAGAGACTGCACAGACACAACTCACTACTGTATGT 5002
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 GGCACTGTCTTGGGCCCTCCTGTGATAGAGACTGCACAGACACAACTCACTACTGTATGT 3000

Qy       5003 TTGTAAAACATCTTAATTTGGTTCTCTAGACTGCTACAAACAAAGGTGCTGCCCAGTCAT 5062
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 TTGTAAAACATCTTAATTTGGTTCTCTAGACTGCTACAAACAAAGGTGCTGCCCAGTCAT 3060

Qy       5063 GTCAAGAGGGATAA 5076
              ||||||||||||||
Db       3061 GTCAAGAGGGATAA 3074


RESULT 11
AX235411
LOCUS       AX235411                5107 bp    DNA     linear   PAT 11-SEP-2001
DEFINITION  Sequence 58 from Patent WO0162928.
ACCESSION   AX235411
VERSION     AX235411.1  GI:15546025
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 5107)
  AUTHORS   Vernet,C.A., Fernandes,E., Shimkets,R.A., Macdougall,J. and
            Spaderna,S.K.
  TITLE     Polypeptides and nucleic acids encoding same
  JOURNAL   Patent: WO 0162928-A 58 30-AUG-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1..5107
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT     1445 a   1061 c   1235 g   1366 t
ORIGIN

  Query Match             60.5%;  Score 3070.8;  DB 6;  Length 5107;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 3072;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2003 TAGCAGTGTGCTTACATATCCAGACCCAGCAGACAGTCAATGACAGCTTGTGTGATATGG 2062
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TAGCAGTGTGCTTACATATCCAGACCCAGCAGACAGTCAATGACAGCTTGTGTGATATGG 60

Qy       2063 TCCACCGTCCTCCAGCCATGAGCCAGGCCTGTAACACAGAGCCCCTGCCCCCCAGGTGGC 2122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCCACCGTCCTCCAGCCATGAGCCAGGCCTGTAACACAGAGCCCCTGCCCCCCAGGTGGC 120

Qy       2123 ATGTGGGACTCTTGGGGGCCTGCTCAGCTACCTGTGGAGTTGGAATTCAGACCCGAGATG 2182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ATGTGGGACTCTTGGGGGCCTGCTCAGCTACCTGTGGAGTTGGAATTCAGACCCGAGATG 180

Qy       2183 TGTACTGCCTGCACCCAGGGGAGACCCTGCCCCTCCTGAGGAGCCCTCGAGATGAAAAGC 2242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGTACTGCCTGCACCCAGGGGAGACCCTGCCCCTCCTGAGGAGCCCTCGAGATGAAAAGC 240

Qy       2243 CCCATGCTTTACAAGCATGCAATCAGTTTGACTGCCCTCCTGGCTGGCACATTGAAGAAT 2302
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCCATGCTTTACAAGCATGCAATCAGTTTGACTGCCCTCCTGGCTGGCACATTGAAGAAT 300

Qy       2303 GGAAGCAGTGTTCCAGGACTCGGGCTGGGGAACTCAGAACAGAAGAGTCACCCTGTCGGC 2362
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGCAGCAGTGTTCCAGGACTTGTGGCGGGGGAACTCAGAACAGAAGAGTCACCTGTCGGC 360

Qy       2363 AGCTGCTAACGGATGGCAGCTTTTTGAATCTCTCAGATGAATTGTGCCAAGGACCCAAGG 2422
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AGCTGCTAACGGATGGCAGCTTTTTGAATCTCTCAGATGAATTGTGCCAAGGACCCAAGG 420

Qy       2423 CATCGTCTCACAAGTCCTGTGCCAGGACAGACTGTCCTCCACATTTAGCTGTGGGAGACT 2482
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CATCGTCTCACAAGTCCTGTGCCAGGACAGACTGTCCTCCACATTTAGCTGTGGGAGACT 480

Qy       2483 GGTCGAAGTGTTCTGTCAGTTGTGGTGTTGGAATCCAGAGAAGAAAGCAGGTGTGTCAAA 2542
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GGTCGAAGTGTTCTGTCAGTTGTGGTGTTGGAATCCAGAGAAGAAAGCAGGTGTGTCAAA 540

Qy       2543 GGCTGGCAGCCAAAGGTCGGCGCATCCCCCTCAGTGAGATGATGTGCAGGGATCTACCAG 2602
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GGCTGGCAGCCAAAGGTCGGCGCATCCCCCTCAGTGAGATGATGTGCAGGGATCTACCAG 600

Qy       2603 GGTTCCCTCTTGTAAGATCTTGCCAGATGCCTGAGTGCAGTAAAATCAAATCAGAGATGA 2662
              || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GGCTCCCTCTTGTAAGATCTTGCCAGATGCCTGAGTGCAGTAAAATCAAATCAGAGATGA 660

Qy       2663 AGACAAAACTTGGTGAGCAGGGTCCGCAGATCCTCAGTGTCCAGAGAGTCTACATTCAGA 2722
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AGACAAAACTTGGTGAGCAGGGTCCGCAGATCCTCAGTGTCCAGAGAGTCTACATTCAGA 720

Qy       2723 CAAGGGAAGAGAAGCGTATTAACCTGACCATTGGTAGCAGAGCCTATTTGCTGCCCAACA 2782
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CAAGGGAAGAGAAGCGTATTAACCTGACCATTGGTAGCAGAGCCTATTTGCTGCCCAACA 780

Qy       2783 CATCCGTGATTATTAAGTGCCCCGTGCCGACGATTCCAGAAATCTCTGATCCAGTGGAGA 2842
              ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        781 CATCCGTGATTATTAAGTGCCCAGTGCCGACGATTCCAGAAATCTCTGATCCAGTGGAGA 840

Qy       2843 AGGATGGCCGTTGCCTGCAGAACTCCAAACGGCTTGGCATCACCAAGTCAGGCTCACTAA 2902
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AGGATGGCCGTTGCCTGCAGAACTCCAAACGGCTTGGCATCACCAAGTCAGGCTCACTAA 900

Qy       2903 AAATCCACGGTCTTGCTGCCCCGACATCGGCCGTGTACCGGTGCATTGCAGGCTCTGCAC 2962
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AAATCCACGGTCTTGCTGCCCCGACATCGGCCGTGTACCGGTGCATTGCAGGCTCTGCAC 960

Qy       2963 AGGAAACAGTTGTGCTCAAGCTCATTGGTACTGACAACCGGCTCATCGCACGCCCAGCCC 3022
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 AGGAAACAGTTGTGCTCAAGCTCATTGGTACTGACAACCGGCTCATCGCACGCCCAGCCC 1020

Qy       3023 TCAGGGAGCCTATGAGGGAATATCCTGGGATGGACCACAGCGAAGCCAATAGTTTGGGAG 3082
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TCAGGGAGCCTATGAGGGAATATCCTGGGATGGACCACAGCGAAGCCAATACTTTGGGAG 1080

Qy       3083 TCACATGGCACAAAATGAGGCAAATGTGGAATAACAAAAATGACCTTTATCTGGATGATG 3142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TCACATGGCACAAAATGAGGCAAATGTGGAATAACAAAAATGACCTTTATCTGGATGATG 1140

Qy       3143 ACCACATTAGTAACCAGCCTTTCTTGAGAGCTCTGTTAGGCCACTGCAGCAATTCTGCAG 3202
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 ACCACATTAGTAACCAGCCTTTCTTGAGAGCTCTGTTAGGCCACTGCAGCAATTCTGCAG 1200

Qy       3203 GAAGCACCAACTCCTGGGAGTTGAAGAATAAGCAGTTTGAAGCAGCAGTTAAACAAGGAG 3262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GAAGCACCAACTCCTGGGAGTTGAAGAATAAGCAGTTTGAAGCAGCAGTTAAACAAGGAG 1260

Qy       3263 CATATAGCATGGATACAGCCCAGTTTGATGAGCTGATAAGAAACATGAGTCAGCTCATGG 3322
              |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db       1261 CATATAGCATGGATACAGCCCAGTTTGATGAGCTGA.AAGAAACATGAGTCAGCTCATGG 1320

Qy       3323 AAACCGGAGAGGTCAGCGATGATCTTGCGTCCCAGC.GATATATCAGCTGGTGGCCGAAT 3382
              |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db       1321 AAACCGGAGAGGTCAGCGATGATCTTGCGTCCCAGCTGATATATCAGCTGGTGGCCGAAT 1380

Qy       3383 TAGCCAAGGCACAGCCAACACACATGCAGTGGCGGGGCATCCAGGAAGAGACACCTCCTG 3442
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 TAGCCAAGGCACAGCCAACACACATGCAGTGGCGGGGCATCCAGGAAGAGACACCTCCTG 1440

Qy       3443 CTGCTCAGCTCAGAGGGGAAACAGGGAGTGTGTCCCAAAGCTCGCATGCAAAAAACTCAG 3502
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 CTGCTCAGCTCAGAGGGGAAACAGGGAGTGTGTCCCAAAGCTCGCATGCAAAAAACTCAG 1500

Qy       3503 GCAAGCTGACATTCAAGCCGAAAGGACCTGTTCTCATGAGGCAAAGCCAACCTCCCTCAA 3562
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 GCAAGCTGACATTCAAGCCGAAAGGACCTGTTCTCATGAGGCAAAGCCAACCTCCCTCAA 1560

Qy       3563 TTTCATTTAATAAAACAATAAATTCCAGGATTGGAAATACAGTATACATTACAAAAAGGA 3622
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 TTTCATTTAATAAAACAATAAATTCCAGGATTGGAAATACAGTATACATTACAAAAAGGA 1620

Qy       3623 CAGAGGTCATCAATATACTGTGTGACCTTATTACCCCCAGTGAGGCCACATATACATGGA 3682
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 CAGAGGTCATCAATATACTGTGTGACCTTATTACCCCCAGTGAGGCCACATATACATGGA 1680

Qy       3683 CCAAGGATGGAACCTTGTTACAGCCCTCAGTAAAAATAATTTTGGATGGAACTGGGAAGA 3742
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 CCAAGGATGGAACCTTGTTACAGCCCTCAGTAAAAATAATTTTGGATGGAACTGGGAAGA 1740

Qy       3743 TACAGATACAGAATCCTACAAGGAAAGAACAAGGCATATATGAATGTTCTGTAGCTAATC 3802
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 TACAGATACAGAATCCTACAAGGAAAGAACAAGGCATATATGAATGTTCTGTAGCTAATC 1800

Qy       3803 ATCTTGGTTCAGATGTGGAAAGTTCTTCTGTGCTGTATGCAGAGGCACCTGTCATCTTGT 3862
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 ATCTTGGTTCAGATGTGGAAAGTTCTTCTGTGCTGTATGCAGAGGCACCTGTCATCTTGT 1860

Qy       3863 CTGTTGAAAGAAATATCACCAAACCAGAGCACAACCATCTGTCTGTTGTGGTTGGAGGCA 3922
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 CTGTTGAAAGAAATATCACCAAACCAGAGCACAACCATCTGTCTGTTGTGGTTGGAGGCA 1920

Qy       3923 TCGTGGAGGCAGCCCTTGGAGCAAACGTGACAATCCGATGTCCTGTAAAAGGTGTCCCTC 3982
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 TCGTGGAGGCAGCCCTTGGAGCAAACGTGACAATCCGATGTCCTGTAAAAGGTGTCCCTC 1980

Qy       3983 AGCCTAATATAACTTGGTTGAAGAGAGGAGGATCTCTGAGTGGCAATGTTTCCTTGCTTT 4042
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 AGCCTAATATAACTTGGTTGAAGAGAGGAGGATCTCTGAGTGGCAATGTTTCCTTGCTTT 2040

Qy       4043 TCAATGGATCCCTGTTGTTGCAGAATGTTTCCCTTGAAAATGAAGGAACCTACGTCTGCA 4102
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 TCAATGGATCCCTGTTGTTGCAGAATGTTTCCCTTGAAAATGAAGGAACCTACGTCTGCA 2100

Qy       4103 TAGCCACCAATGCTCTTGGAAAGGCAGTGGCAACATCTGTACTCCACTTGCTGGAACGAA 4162
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 TAGCCACCAATGCTCTTGGAAAGGCAGTGGCAACATCTGTACTCCACTTGCTGGAACGAA 2160

Qy       4163 GATGGCCAGAGAGTAGAATCGTATTTCTGCAAGGACATAAAAAGTACATTCTCCAGGCAA 4222
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 GATGGCCAGAGAGTAGAATCGTATTTCTGCAAGGACATAAAAAGTACATTCTCCAGGCAA 2220

Qy       4223 CCAACACTAGAACCAACAGCAATGACCCAACACGAGAACCCCCGCCTCAAGAGCCTTTTT 4282
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 CCAACACTAGAACCAACAGCAATGACCCAACACGAGAACCCCCGCCTCAAGAGCCTTTTT 2280

Qy       4283 GGGAGCCTGGTAACTGGTCACATTGTTCTGCCACCTGTGGTCATTTGGGAGCCCGCATTC 4342
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 GGGAGCCTGGTAACTGGTCACATTGTTCTGCCACCTGTGGTCATTTGGGAGCCCGCATTC 2340

Qy       4343 AGAGACCCCAGTGTGTGATGGCCAATGGGCAGGAAGTGAGTGAGGCCCTGTGTGATCACC 4402
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 AGAGACCCCAGTGTGTGATGGCCAATGGGCAGGAAGTGAGTGAGGCCCTGTGTGATCACC 2400

Qy       4403 TCCAGAAGCCACTGGCTGGGTTTGAGCCCTGTAACATCCGGGACTGCCCAGCGAGGTGGT 4462
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 TCCAGAAGCCACTGGCTGGGTTTGAGCCCTGTAACATCCGGGACTGCCCAGCGAGGTGGT 2460

Qy       4463 TCACAAGTGTGTGGTCACAGTGCTCTGTGTCTTGCGGTGAAGGATACCACAGTCGGCAGG 4522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 TCACAAGTCTGTGGTCACAGTGCTCTGTGTCTTGCGGTGAAGGATACCACAGTCGGCAGG 2520
```

```
Qy       4523 TGACGTGTAAGGGGACAAAAGTCAATGGAACTGTGCAGGTGGTGTCTCCAAGAGCATGTG 4582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 TGACGTGTAAGCGGACAAAAGCCAATGGAACTGTGCAGGTGGTGTCTCCAAGAGCATGTG 2580

Qy       4583 CCCCTAAAGACCGGCCTCTGGGAAGAAAACCATGTTTTGGTCATCCATGTGTTCAGTGGG 4642
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 CCCCTAAAGACCGGCCTCTGGGAAGAAAACCATGTTTTGGTCATCCATGTGTTCAGTGGG 2640

Qy       4643 AACCAGGGAACCGGTGTCCTGGACGTTGCATGGGCCGTGCTGTGAGGATGCAGCAGCGTC 4702
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 AACCAGGGAACCGGTGTCCTGGACGTTGCATGGGCCGTGCTGTGAGGATGCAGCAGCGTC 2700

Qy       4703 ACACAGCTTGTCAACACAACAGCTCTGACTCCAACTGTGATGACAGAAAGAGACCCACCT 4762
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 ACACAGCTTGTCAACACAACAGCTCTGACTCCAACTGTGATGACAGAAAGAGACCCACCT 2760

Qy       4763 TAAGAAGGAACTGCACATCAGGGGCCTGTGATGTGTGTTGGCACACAGGCCCTTGGAAGC 4822
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 TAAGAAGGAACTGCACATCAGGGGCCTGTGATGTGTGTTGGCACACAGGCCCTTGGAAGC 2820

Qy       4823 CCTGTACAGCAGCCTGTGGCAGGGGTTTCCAGTCTCGGAAAGTCGACTGTATCCACACAA 4882
              |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db       2821 CCTGTACAGCAGCCTGTGGCAGGGCTTTCCAGTCTCGGAAAGTCGACTGTATCCACACAA 2880

Qy       4883 GGAGTTGCAAACCTGTGGGCAAGAGACACTGTGTACAGAAAAAGAAACCAATTTCCTGGC 4942
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 GGAGTTGCAAACCTGTGGGCAAGAGACACTGTGTACAGAAAAAGAAACCAATTTCCTGGC 2940

Qy       4943 GGCACTGTCTTGGGCCCTCCTGTGATAGAGACTGCACAGACACAACTCACTACTGTATGT 5002
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 GGCACTGTCTTGGGCCCTCCTGTGATAGAGACTGCACAGACACAACTCACTACTGTATGT 3000

Qy       5003 TTGTAAAACATCTTAATTTGTGTTCTCTAGACCGCTACAAACAAAGGTGCTGCCAGTCAT 5062
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 TTGTAAAACATCTTAATTTGTGTTCTCTAGACCGCTACAAACAAAGGTGCTGCCAGTCAT 3060

Qy       5063 GTCAAGAGGGATAA 5076
              ||||||||||||||
Db       3061 GTCAAGAGGGATAA 3074


RESULT 12
AX149471
LOCUS       AX149471                2912 bp    DNA     linear   PAT 08-JUN-2001
DEFINITION  Sequence 27 from Patent WO0136638.
ACCESSION   AX149471
VERSION     AX149471.1  GI:14347941
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2912)
  AUTHORS   Shimkets,R.A., Lichenstein,H., Vernet,C. and Fernandes,E
  TITLE     Polypeptides and nucleic acids encoding same
  JOURNAL   Patent: WO 0136638-A 27 25-MAY-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1..2912
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      813 a    710 c    754 g    635 t
ORIGIN

  Query Match             56.4%;  Score 2862.8;  DB 6;  Length 2912;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 2867;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy        843 CAGCTTTAACAGCCCGGGGTCTTTGTCGTAGAAAAACAACAGTGGAATTTTAGAGGGG 902
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAGCTTTAACAGCCCGGGGTCTTTGTCGTAGAAAAACAACAGTGGAATTTTAGAGGGG 60

Qy        903 CTCCGAGAGGAAACTTTTAAGATCCCAGGACCTCTGATGGCTGATTTCATCTTCAAGAC 962
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTCCGAGAGGAAACTTTTAAGATCCCAGGCCCTTGATGGCTGATTTCATCTTCAAGAC 120

Qy        963 CAGGTACACTGCAGCCAAAGACAGCCGTGGTTCAGTTCTTCTTTTACCAGCCCATCAGTCA 1022
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CAGGTACACTGCAGCCAAAGACAGCCGTGGTTCAGTTCTTCTTTTACCAGCCCATCAGTCA 180

Qy       1023 TCAGTGGAGACAAACTGACTTCTTTCCCTGCACTCTGACGTGTGGAGGAGGTTATCAGCT 1082
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TCAGTGGAGACAAACTGACTTCTTTCCCTGCACTCTGACGTGTGGAGGAGGTTATCAGCT 240

Qy       1083 CAATTCTGCTGAATGTGTGGATATCCGCTTGAAGAGGGTAGTTCCTGACCATTATTGTCA 1142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CAATTCTGCTGAATGTGTGGATATCCGCTTGAAGAGGGTAGTTCCTGACCATTATTGTCA 300

Qy       1143 CTACTACCCTGAAAATGTAAAACCAAAACCAAAACTGAAGGAATGCAGCATGGATCCCTG 1202
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CTACTACCCTGAAAATGTAAAACCAAAACCAAAACTGAAGGAATGCAGCATGGATCCCTG 360

Qy       1203 CCCATCAAGTGATGGATTTAAAGAGATAATGCCCTATGACCACTTCCAACCTCTTCCTCG 1262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCCATCAAGTGATGGATTTAAAGAGATAATGCCCTATGACCACTTCCAACCTCTTCCTCG 420

Qy       1263 CTGGGAACATAATCCTTGGACTGCATGTTCCGTGTCCTGTGGAGGAGGGATTCAGAGACG 1322
              |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        421 CTGGGAACATAATCCTTGGACTGCATGTTCCGTGTCCTGTGGAGGAGGGATTCAGAGACG 480

Qy       1323 GAGCTTTGTGTGTGTAGAGGAATCCATGCATGGAGAGATATTGCAGGTGGAAGAATGGAA 1382
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAGCTTTGTGTGTGTAGAGGAATCCATGCATGGAGAGATATTGCAGGTGGAAGAATGGAA 540

Qy       1383 GTGCATGTACGCACCCAAACCCAAGGTTATGCAAACTTGTAATCTGTTTGATTGCCCCAA 1442
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        541 GTGCATGTACGCACCCAAACCCAAGGTTATGCAAACTTGTAATCTGTTTGATTGCCCCAA 600

Qy       1443 GTGGATTGCCATGGAGTGGTCTCAGTGCACAGTGACTTGTGGCCGAGGGTTACCGTACCG 1502
              |||||||||||||||||||||||||||||||||||| ||||||||||||||| |||||
Db        601 GTGGATTGCCATGGAGTGGTCTCAGTGCACAGTGACTTGTGGCCGAGGGTTACCGTACCG 660

Qy       1503 GGTTGTTCTGTGTATTAACCACCGCGGAGAGCATGTTGGGGGCTGCAATCCACAACTGAA 1562
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GGTTGTTCTGTGTATTAACCACCGCGGAGAGCATGTTGGGGGCTGCAATCCACAACTGAA 720

Qy       1563 GTTACACATCAAAGAAGAATGTGTCATTCCCATCCCGTGTTATAAACCAAAAGAAAAAAG 1622
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GTTACACATCAAAGAAGAATGTGTCATTCCCATCCCGTGTTATAAACCAAAAGAAAAAAG 780

Qy       1623 TCCAGTGGAAGCAAAATTGGCTTGGCTGAAACAAGCACAAGAACTAGAAGAGACCAGAAT 1682
              ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        781 TCCAGTGGAAGCAAAATTGGCTTGGCTGAAACAAGCACAAGAACTAGAAGAGACCAGAAT 840

Qy       1683 AGCAACAGAAGAACCAACGTTCATTCCAGAACCCTGGTCAGCCTGCAGTACCACCGTGGG 1742
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AGCAACAGAAGAACCAACGTTCATTCCAGAACCCTGGTCAGCCTGCAGTACCACCGTGGG 900

Qy       1743 GCCAGGTGTGCAGGTCCGCGAGGTGAAGTGCCGTGTGCTCCTCACATTCACGCAGACTGA 1802
              ||| |||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db        901 GCCGGGTGTGCAGGTCCGTGAGGTGAAGTGCCGTGTGCTCCTCACATTCACGCAGACTGA 960

Qy       1803 GACTGAGCTGCCCGAGGAAGAGTGTGAAGGCCCCAAGCTGCCCACCGAACGGCCCTGCCT 1862
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GACTGAGCTGCCCGAGGAAGAGTGTGAAGGCCCCAAGCTGCCCACCGAACGGCCCTGCCT 1020

Qy       1863 CCTGGAAGCATGTGATGAGAGCCCGGCCTCCCGAGAGCTAGACATCCCTCTCCCTGAGGA 1922
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CCTGGAAGCATGTGATGAGAGCCCGGCCTCCCGAGAGCTAGACATCCCTCTCCCTGAGGA 1080

Qy       1923 CAGTGAGACGACTTACGACTGGGAGTACGCTGGGTTCACCCCTTGCACAGCAACATGCTT 1982
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db       1081 CAGTGAGACGACTTACGACTGGGAGTACGCTGGGTTCACCCCTTGCACAGCAACATGCGT 1140

Qy       1983 GGGAGGGCATCAAGAAGCCATAGCAGTGTGCTTACATATCCAGACCTAGCAGACAGTCAA 2042
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 GGGAGGCCATCAAGAAGCCATAGCAGTGTGCTTACATATCCAGACCCAGCAGACAGTCAA 1200

Qy    2043 TGACAGCTTGTGTGATATGGTCCACCGTCCTCCAGCCATGAGCCAGGCCTGTAACACAGA 2102
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 TGACAGCTTGTGTGATATGGTCCACCGTCCTCCAGCCATGAGCCAGGCCTGTAACACAGA 1260

Qy    2103 GCCCTGTCCCCCCAGGTGGCATGTGGGCTCTTGGGGGCCCTGCTCAGCTACCTGTGGAGT 2162
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 GCCCTGTCCCCCCAGGTGGCATGTGGGCTCTTGGGGGCCCTGCTCAGCTACCTGTGGAGT 1320

Qy    2163 TGCAATTCAGACCCCAGATGTGTACTGCCTGCACCCAGGCACACCCCTGCCCCTCCTGA 2222
           |||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db    1321 TGGAATTCAGACCCGAGATGTGTACTGCCTGCACCCAGGGAGACCCCTGCCCCTCCTGA 1380

Qy    2223 GGAGTGCCCAGATGAAAAGCCCCATGCTTTACAAGCATGCAATCAGTTTGACTGCCCTCC 2282
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1381 GGAGTGCCGAGATGAAAAGCCCCATGCTTTACAAGCATGCAATCAGTTTGACTGCCCTCC 1440

Qy    2283 TGGCTGGCACATTGAAGAATGGCAGCAGTGTTCCAGGACTTGTGGCGGGGGAACTCAGAA 2342
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1441 TGGCTGGCACATTGAAGAATGGCAGCAGTGTTCCAGGACTTGTGGCGGGGGAACTCAGAA 1500

Qy    2343 CAGAAGAGTCACCTGTCGGCAGCTGCTAACGGATGGCAGCTTTTTGAATCTCTCAGATGA 2402
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1501 CAGAAGAGTCACCTGTCGGCAGCTGCTAACGGATGGCAGCTTTTTGAATCTCTCAGATGA 1560

Qy    2403 ATTGTGCCAAGGACCCAAGGCATCGTCTCACAAGTCCTGTGCCAGGACAGACTGTCCTCC 2462
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1561 ATTGTGCCAAGGACCCAAGGCATCGTCTCACAAGTCCTGTGCCAGGACAGACTGTCCTCC 1620

Qy    2463 ACATTTAGCTGTGGGAGACTGGTCCAACTCTTCTCTCACTTGTCCTCTTGGAATCCAGAG 2522
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1621 ACATTTAGCTGTGGGAGACTGGTCGAAGTGTTCTGTCAGTTGTGGTGTTGGAATCCAGAG 1680

Qy    2523 AAGAAAGCAGGTGTGTCAAAGGCTGGCAGCCAAAGGTCGGCGCATCCCCCTCAGTGAGAT 2582
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1681 AAGAAAGCAGGTGTGTCAAAGGCTGGCAGCCAAAGGTCGGCGCATCCCCCTCAGTGAGAT 1740

Qy    2583 GATGTGCAGGGATCTACCAGGGTTCCCTCTTGTAAGATCTTGCCAGATGCCTGACTGCAG 2642
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1741 GATGTGCAGGGATCTACCAGGGTTCCCTCTTGTAAGATCTTGCCAGATGCCTGAGTGCAG 1800

Qy    2643 TAAAATCAAATCAGAGATGAAGACAAAACTTGGTGAGCAGGGTCCGCAGATCCTCAGTGT 2702
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1801 TAAAATCAAATCAGAGATGAAGACAAAACTTGGTGAGCAGGGTCCGCAGATCCTCAGTGT 1860

Qy    2703 CCAGAGAGTCTACATTCAGACAAGGGAAGAGAAGCGTATTAACCTCACCATTGGTAGCAG 2762
           ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db    1861 CCAGAGAGTCTACATTCAGACAAGGGAAGAGAAGCGTATTAACCTCACCATTGGTAGCAG 1920

Qy    2763 AGCCTATTTGCTGCCCAACACATCCGTCATTATTAACTGCCCCGTGCGACGATTCCAGAA 2822
           ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db    1921 AGCCTATTTGCTGCCCAACACATCCGTGATTATTAAGTGCCCAGTGCGACGATTCCAGAA 1980

Qy    2823 ATCTCTGATCCACTGGCAGAAGGATCGCCGTTGCCTGCAGAACTCCAAACGGCTTGGCAT 2882
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1981 ATCTCTGATCCACTGGCAGAAGGATCGCCGTTGCCTGCAGAACTCCAAACGGCTTGGCAT 2040

Qy    2883 CACCAAGTCAGGCTCACTAAAAATCCACGGTCTTGCTGCCCCCCACATCGGCGTGTACCG 2942
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2041 CACCAAGTCAGGCTCACTAAAAATCCACGGTCTTGCTGCCCCCGACATCGGCGTGTACCG 2100

Qy    2943 GTGCATTGCAGGCTCTGCACAGGAAACAGTTGTGCTCAAGCTCATTGGTACTGACAACCG 3002
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2101 GTGCATTGCAGGCTCTGCACAGGAAACAGTTGTGCTCAAGCTCATTGGTACTGACAACCG 2160

Qy    3003 GCTCATCGCACGGCCCAGCCCTCAGGGAGCCTATGAGGGAATATCCTGGGATGGACCACAG 3062
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2161 GCTCATCGCACGGCCCACCCCTCAGGGAGCCTATGAGGGAATATCCTGGGATGGACCACAG 2220

Qy    3063 CGAAGCCAATAGTTTGGGAGTCACATGGCACAAAATGAGGCAAATGTGGAATAACAAAAA 3122
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2221 CCAACCCAATAGTTTGGCAGTCACATGGCACAAAATGAGGCAAATGTGGAATAACAAAAA 2280

Qy    3123 TGACCTTTATCTGGATGATGACCACATTAGTAACCAGCCTTTCTTGAGAGCTCTGTTAGG 3182
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2281 TGACCTTTATCTGGATGATGACCACATTAGTAACCAGCCTTTCTTGAGAGCTCTGTTAGG 2340

Qy    3183 CCACTGCAGCAATTCTGCAGGAAGCACCAACTCCTGGGAGTTGAAGAATAAGCAGTTTGA 3242
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2341 CCACTGCAGCAATTCTGCAGGAAGCACCAACTCCTGGGAGTTGAAGAATAAGCAGTTTGA 2400

Qy    3243 AGCAGCAGTTAAACAAGGAGCATATAGCATGGATACAGCCCAGTTTGATGAGCTGATAAG 3302
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2401 AGCAGCAGTTAAACAAGGAGCATATAGCATGGATACAGCCCAGTTTGATGAGCTGATAAG 2460

Qy    3303 AAACATGAGTCAGCTCATGGAAACCGGAGAGGTCAGCGATGATCTTGCGTCCCAGCTGAT 3362
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2461 AAACATGAGTCAGCTCATGGAAACCGGAGAGGTCAGCGATGATCTTGCGTCCCAGCTGAT 2520

Qy    3363 ATATCAGCTGGTGGCCGAATTAGCCAAGGCACAGCCAACACACATGCAGTGGCGGGGCAT 3422
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2521 ATATCAGCTGGTGGCCGAATTAGCCAAGGCACAGCCAACACACATGCAGTGGCGGGGCAT 2580

Qy    3423 CCAGGAAGAGACACCTCCTGCTGCTCAGCTCAGAGGGAAACAGGGAGTGTGTCCCAAAG 3482
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2581 CCAGGAAGAGACACCTCCTGCTGCTCAGCTCAGAGGGAAACAGGGAGTGTGTCCCAAAG 2640

Qy    3483 CTCGCATGCAAAAAACTCAGGCAAGCTGACATTCAAGCCGAAAGGACCTGTTCTCATGAG 3542
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2641 CTCGCATGCAAAAAACTCAGGCAAGCTGACATTCAAGCCGAAAGGACCTGTTCTCATGAG 2700

Qy    3543 GCAAAGCCAACCTCCCTCAATTTCATTTAATAAAACAATAAATTCCAGGATTGGAAATAC 3602
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2701 GCAAAGCCAACCTCCCTCAATTTCATTTAATAAAACAATAAATTCCAGGATTGGAAATAC 2760

Qy    3603 AGTATACATTACAAAAAGGACAGAGGTCATCAATATACTGTGTGACCTTATTACCCCCAG 3662
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2761 AGTATACATTACAAAAAGGACAGAGGTCATCAATATACTGTGTGACCTTATTACCCCCAG 2820

Qy    3663 TGAGGCCACATATACATGGACCAAGGATGGAACCTTGTTACAGCCCTCAGTAAA 3716
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2821 TCAGCCCACATATACATGGACCAAGGATGGAACCTTCTTACAGCCCTCAGTAAA 2874


RESULT 13
AX224819
LOCUS       AX224819                2538 bp    DNA     linear   PAT 10 SEP 2001
DEFINITION  Sequence 11 from Patent WO0161011.
ACCESSION   AX224819
VERSION     AX224819.1  GI:15554914
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2538)
  AUTHORS   Donoho,G., Scoville,J., Turner,C.A., Friedrich,G., Zambrowicz,B.
            and Sands,A.T.
  TITLE     Novel human thrombospondin repeat proteins and polynucleotides
            encoding the same
  JOURNAL   Patent: WO 0161011-A 11 23 AUG 2001;
            Lexicon Genetics Incorporated (US)
FEATURES             Location/Qualifiers
     source          1..2538
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      670 a    627 c    678 g    563 t
ORIGIN

  Query Match             49.1%;  Score 2490;  DB 6;  Length 2538;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 2490;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 ATGGCTTCCTGGACGAGCCCCTGGTGGGTGCTGATAGGGATGGTCTTCATGCACTCTCCC 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGCTTCCTGGAGGAGCCCCTGGTGGGTGCTGATAGGGATGGTCTTCATGCAGCTCTCC 60

Qy         61 CTGCCGCAGACCACAGCTGAGAAATCTCCTGGAGCCTATTTCCTTCCCGAGTTTGCATTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGCCGCAGACCACAGCTGAGAAATCTCCTGGAGCCTATTTCCTTCCCGAGTTTGCATTT 120

Qy        121 TCTCCTCAGGGAAGTTTTCTGGAAGACACAACAGGGGAGCAGTTCCTCACTTATCGCTAT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCTCCTCAGGGAAGTTTTCTGGAAGACACAACAGGGGAGCAGTTCCTCACTTATCGCTAT 180

Qy        181 GATGACCAGACCTCAAGAAACACTCGTTCAGATGAAGACAAAGATGGCAACTGGGATGCT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GATGACCAGACCTCAAGAAACACTCGTTCAGATGAAGACAAAGATGGCAACTGGGATGCT 240

Qy        241 TGGGGCGACTGGAGTGACTGGTCCCGGACCTGTGGGGGAGGAGCATCATATTCCCTGCGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGGGGCGACTGGAGTGACTGGTCCCGGACCTGTGGGGGAGGAGCATCATATTCCCTGCGG 300

Qy        301 AGATCTTTGACTGGAAGGAATTCTGAAGGGAGAAACATTCGGTACAAGACATGCAGCAAT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGATCTTTGACTGGAAGGAATTCTGAAGGGAGAAACATTCGGTACAAGACATGCAGCAAT 360

Qy        361 CATGACTGCCCTCCAGATGCAGAAGATTTCAGAGCCTAGCAGTGGCTAGCCTACAATGAT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CATGACTGCCCTCCAGATGCAGAAGATTTCAGAGCCTAGCAGTGGCTAGCCTACAATGAT 420

Qy        421 GTCCAGTATCAGGGCATTACTATGAATGGCTTCCACGATATAATGATCCTGCTGCCCCCG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GTCCAGTATCAGGGCATTACTATGAATGGCTTCCACGATATAATGATCCTGCTGCCCCCG 480

Qy        481 TGTGCACTCAAGTGTCATGCACAAGGACAAAACTTGGTGGTGGAGCTGGCACCTAAGGTA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TGTGCACTCAAGTGTCATGCACAAGGACAAAACTTGGTGGTGGAGCTGGCACCTAAGGTA 540

Qy        541 CTGGATGGAACTCGTTGCAACACGGACTCCTTGGACATGTGTATCAGTGGCATCTGTCAG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTGGATGGAACTCGTTGCAACACGGACTCCTTGGACATGTGTATCAGTGGCATCTGTCAG 600

Qy        601 GCAGTGGGCTGCGATCGGCAACTGGGAAGCAATGCCAAGGAGGACAACTGTGGAGTCTGT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GCAGTGGGCTGCGATCGGCAACTGGGAAGCAATGCCAAGGAGGACAACTGTGGAGTCTGT 660

Qy        661 GCCGGGGATGGCTCCACCTGCAGGCTTGTACGGGGACAATCAAAGTCACACGTTTCTCCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GCCGGGGATGGCTCCACCTGCAGGCTTGTACGGGGACAATCAAAGTCACACGTTTCTCCT 720

Qy        721 GAAAAAAGAGAAGAAAATGTAATGCCTGTTTCCTTGGGAAGTCGAAGTGTGAGAATTACA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GAAAAAAGAGAAGAAAATGTAATGCCTGTTTCCTTGGGAAGTCGAAGTGTGAGAATTACA 780

Qy        781 GTGAAAGGACCTGCCCACCTCTTTATTGAATCAAAAACACTTCAAGGAAGCAAAGGAGAA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GTGAAAGGACCTGCCCACCTCTTTATTGAATCAAAAACACTTCAAGGAAGCAAAGGAGAA 840

Qy        841 CACAGCTTTAACAGCCCCGGCGTCTTTGTCGTAGAAAACACAACAGTGGAATTTCAGAGG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CACAGCTTTAACAGCCCCGGCGTCTTTGTCGTAGAAAACACAACAGTGGAATTTCAGAGG 900

Qy        901 GGCTCCGAGAGGCAAACTTTTAAGATTCCAGGACCTCTGATGGCTGATTTCATCTTCAAG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GGCTCCGAGAGGCAAACTTTTAAGATTCCAGGACCTCTGATGGCTGATTTCATCTTCAAG 960

Qy        961 ACCAGGTACACTGCAGCCAAAGACAGCGTGGTTCAGTTCTTCTTTTACCAGCCCATCAGC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 ACCAGGTACACTGCAGCCAAAGACAGCGTGGTTCAGTTCTTCTTTTACCAGCCCATCAGC 1020

Qy       1021 CATCAGTGGAGACAAACTGACTGCTTTCCTCGCACTGTGACGTGTGGAGGAGGTTATCAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CATCAGTGGAGACAAACTGACTGCTTTCCTCGCACTGTGACGTGTGGAGGAGGTTATCAG 1080

Qy       1081 CTCAATTCTGGCTGAATGTGGGATATCCGCTTGAAGAGGGTAGTTCCTGACCATTATTGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CTCAATTCTGGCTGAATGTGGGATATCCGCTTGAAGAGGGTAGTTCCTGACCATTATTGT 1140

Qy       1141 CACTACTACCTGAAAATGTAAAACCAAAACCAAAACTGAAGGAATGCAGCAGCAGGATCC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CACTACTACCTGAAAATGTAAAACCAAAACCAAAACTGAAGGAATGCAGCAGCAGGATCC 1200

Qy       1201 TGCCCATCAAGTGATGGATTTAAAGAGATAATGCCCTATGACCACTTCCAACCTCTTCCT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TGCCCATCAAGTGATGGATTTAAAGAGATAATGCCCTATGACCACTTCCAACCTCTTCCT 1260

Qy       1261 CGCTGGGAACATAATCCTTGGACTGCATGTTCCGTGTCCTGTGGAGGAGGGATTCAGAGA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CGCTGGGAACATAATCCTTGGACTGCATGTTCCGTGTCCTGTGGAGGAGGGATTCAGAGA 1320

Qy       1321 CGGAGCTTTGTGTGTGTAGAGGAATCCATGCATGGAGAGATATTGCAGTTGGAAGAATGG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CGGAGCTTTGTGTGTGTAGAGGAATCCATGCATGGAGAGATATTGCAGTTGGAAGAATGG 1380

Qy       1381 AAGTGCATGTACGCACCCAAAACCAAGGTTATGCAAACTTGTAATCTGTTTGATTGCCCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 AAGTGCATGTACGCACCCAAAACCAAGGTTATGCAAACTTGTAATCTGTTTGATTGCCCC 1440

Qy       1441 AAGTGGATTGCCATGGAGTGGTCTCAGTGCACAGTGACTTGTGGCCGAGGGTTACGGTAC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 AAGTGGATTGCCATGGAGTGGTCTCAGTGCACAGTGACTTGTGGCCGAGGGTTACGGTAC 1500

Qy       1501 CGGGTTGTTCTGTGTATTAACCACCGCGGAGAGCATGTTGGGGGCTGCAATCCACAACTG 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 CGGGTTGTTCTGTGTATTAACCACCGCGGAGAGCATGTTGGGGGCTGCAATCCACAACTG 1560

Qy       1561 AAGTTACACATCAAAGAAGAATGTGTCATTCCCATCCCGTGTTATAAACCAAAAGAAAAA 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 AAGTTACACATCAAAGAAGAATGTGTCATTCCCATCCCGTGTTATAAACCAAAAGAAAAA 1620

Qy       1621 AGTCCAGTGGAAGCAAAATTGCCTTGGCTGAAACAAGCACAAGAACTAGAAGAGACCAGA 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 AGTCCAGTGGAAGCAAAATTGCCTTGGCTGAAACAAGCACAAGAACTAGAAGAGACCAGA 1680

Qy       1681 ATAGCAACAGAAGAACCAACGTTCATTCCAGAACCCTGGTCAGCCTGCAGTACCACGTGT 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 ATAGCAACAGAAGAACCAACGTTCATTCCAGAACCCTGGTCAGCCTGCAGTACCACGTGT 1740

Qy       1741 GGGCCAGGTGTGCAGGTCCGCGAGGTGAAGTGCCGTGTGCTCCTCACATTCACCCAGACT 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 GGGCCAGGTGTGCAGGTCCGCGAGGTGAAGTGCCGTGTGCTCCTCACATTCACCCAGACT 1800

Qy       1801 GAGACTGAGCTGCCCGAGGAAGAGTGTGAAGGCCCCAAGCTGCCCACCGAACGGCCCTGC 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 GAGACTGAGCTGCCCGAGGAAGAGTGTGAAGGCCCCAAGCTGCCCACCGAACGGCCCTGC 1860

Qy       1861 CTCCTGGAAGCATGTGATGAGAGCCCGGCCTCCCGAGAGCTAGACATCCCTCTCCCTGAG 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 CTCCTGGAAGCATGTGATGAGAGCCCGGCCTCCCGAGAGCTAGACATCCCTCTCCCTGAG 1920

Qy       1921 GACAGTGAGACGACTTACGACTGGGAGTACGCTGGGTTCACCCCTTGCACAGCAACATGC 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 GACAGTGAGACGACTTACGACTGGGAGTACGCTGGGTTCACCCCTTGCACAGCAACATGC 1980

Qy       1981 TTGGGAGGCCATCAAGAAGCCATAGCAGTGTGCTTACATATCCAGACCCAGCAGACAGTC 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 TTGGGAGGCCATCAAGAAGCCATAGCAGTGTGCTTACATATCCAGACCCAGCAGACAGTC 2040

Qy       2041 AATGACAGCTTGTGTGATATGGTCCACGTGCTCCACCCATGACCCAGGCCCTGTAACACA 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 AATGACAGCTTGTGTGATATGGTCCACGTGCTCCACCCATGACCCAGGCCCTGTAACACA 2100

Qy       2101 GAGCCTGTCCCCCAGGTGGCATCTGCCCCTCCTTGCCTGGCCCTCCCAGCTACCTCTGGA 2160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 GAGCCTGTCCCCCAGGTGGCATCTGCCCCTCCTTGCCTGGCCCTCCCAGCTACCTCTGGA 2160

Qy       2161 GTTGGAATTCAGACCCGAGATCCTTACTCCTGCACCCAGGGAGACCCCTGCCCCCTCCT 2220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 GTTGGAATTCAGACCCGAGATCCTTACTCCTGCACCCAGGGAGACCCCTGCCCCCTCCT 2220
```

```
Qy     2221 GAGGAGTGCCGAGATGAAAAGCCCCATGCTTTACAAGCATGCAATCAGTTTGACTGCCCT 2280
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2221 GAGGAGTGCCGAGATGAAAAGCCCCATGCTTTACAAGCATGCAATCAGTTTGACTGCCCT 2280

Qy     2281 CCTGGCTGGCACATTGAAGAATGGCAGCAGTGTTCCAGGACTTGTGCCGGGGGAACTCAG 2340
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2281 CCTGGCTGGCACATTGAAGAATGGCAGCAGTGTTCCAGGACTTGTGCCGGGGGAACTCAG 2340

Qy     2341 AACAGAAGAGTCACCTGTCGGCAGCTGCTAACGGATGGCAGCTTTTTGAATCTCTCAGAT 2400
            ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db     2341 AACAGAAGAGTCACCTGTCGGCAGCTGCTAACGGATGGCAGCTTTTTGAATCTCTCAGAT 2400

Qy     2401 GAATTGTGCCAAGGACCCAAGGCATCGTCTCACAAGTCCTGTGCCAGGACAGACTGTCCT 2460
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2401 GAATTGTGCCAAGGACCCAAGGCATCGTCTCACAAGTCCTGTGCCAGGACAGACTGTCCT 2460

Qy     2461 CCACATTTAGCTGTGGGAGACTGGTCGAAG 2490
            ||||||||||||||||||||||||||||||
Db     2461 CCACATTTAGCTGTGGGAGACTGGTCGAAG 2490


RESULT 14
AF237652
LOCUS       AF237652                2523 bp    mRNA    linear   PRI 03-MAR-2001
DEFINITION  Homo sapiens a disintegrin-like and metalloprotease domain with
            thrombospondin type I motifs-like 3 (ADAMTSL3) mRNA, partial cds
ACCESSION   AF237652
VERSION     AF237652.1  GI:13183077
KEYWORDS    .
SOURCE      Homo sapiens
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2523)
  AUTHORS   Hirohata,S., Anand-Apte,B., Seldin,M. and Apte,S.
  TITLE     Punctin, a member of a new family with similarities to ADAM-TS
            proteases, is a component of extracellular matrix of skeletal
            muscle
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2523)
  AUTHORS   Anand-Apte,B. and Apte,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-FEB-2000) Biomedical Engineering, Cleveland Clinic
            Foundation, 9500 Euclid, Cleveland, OH 44195, USA
FEATURES             Location/Qualifiers
     source          1..2523
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     gene            1..>2523
                     /gene="ADAMTSL3"
     CDS             225..>2523
                     /gene="ADAMTSL3"
                     /note="ADAM-TSL3; similar to GenBank Accession Number
                     AB033059; extracellular matrix protein"
                     /codon_start=1
                     /product="a disintegrin like and metalloprotease domain
                     with thrombospondin type I motifs-like 3"
                     /protein_id="AAK15041.1"
                     /db_xref="GI:13183078"
                     /translation="MASWTSPWWVLIGMVFMHSPLPQTTAEKSPGAYFLPEFALSPQG
                     SFLEDTTGEQFLTYRYDDQTSRNTRSDEDKDGNWDAWGDWSDCSRTCGGGASYSLRRC
                     LTGRNCEGQNIRYKTCSNHDCPPDAEDFRAQQCSAYNDVQYQGHYYEWLPRYNDPAAP
                     CALKCHAQGQNLVVELAPKVLDGTRCNTDSLDMCISGICQAVGCDRQLGSNAKEDNCG
                     VCAGDGSTCRLVRGQSKSHVSPEKREENVIAVPLGSRSVRITVKGPAHLFIESKTLQG
                     SKGEHSFNSPGVFVVENTTVEFQRGSERQTFKIPGPLMADFIFKTRYTAAKDSVVQFF
                     FYQPISHQWRQTDFFPCTVTCGGGYQLNSAECVDIRLKRVVPDHYCHYYPENVKPKPK
                     LKECSMDPCPSSDGFKEIMPYDHFQPLPRWEHNPWTACSVSCGGGIQRRSFVCVEESM
                     HGEILQVEEWKCMYAPKPKVMQTCNLFDCPKWIAMEWSQCTVTCGRGLRYRVVLCINH
                     RGEHVGGCNPQLKLHIKEECVIPIPCYKPKEKSPVEAKLPWLKQAQELEETRIATEEP
                     TFIPEPWSACSTTCGPGVQVREVKCRVLLTFTQTETELPEEECEGPKLPTERPCLLEA
                     CDESPASRELDIPLPEDSETTYDWEYAGFTPCTATCLGGHQEAIAVCLHIQTQQTVND
                     SLCDMVHRPPAMSQACNTEPCPPRWHVGSWGPCSATCGVGIQTRDVYCLHPGETPAPP
                     EECRDEKPHALQACNQFDCPPGWHIE"
BASE COUNT      635 a    652 c    692 g    544 t
ORIGIN

  Query Match             45.2%;  Score 2295.8;  DB 9;  Length 2523;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 2297;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy        1 ATGGCTTCCTGGACGAGCCCCTGGTGGGTGCTGATAGGGATGGTCTTCATGCACTCTCCC 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      225 ATGGCTTCCTGGACGAGCCCCTGGTGGGTGCTGATAGGGATGGTCTTCATGCACTCTCCC 284

Qy       61 CTCCCGCAGACCACAGCTGAGAAATCTCCTGGAGCCTATTTCCTTCCCGAGTTTGCACTT 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      285 CTCCCGCAGACCACAGCTGAGAAATCTCCTGGAGCCTATTTCCTTCCCGAGTTTGCACTT 344

Qy      121 TCTCCTCAGGGAAGTTTTCTGGAAGACACAACAGGGGAGCAGTTCCTCACTTATCGCTAT 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      345 TCTCCTCAGGGAAGTTTTCTGGAAGACACAACAGGGGAGCAGTTCCTCACTTATCGCTAT 404

Qy      181 GATGACCAGACCTCAAGAAACACTCGTTCAGATGAAGACAAAGATGGCAACTGGGATGCT 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      405 GATGACCAGACCTCAAGAAACACTCGTTCAGATGAAGACAAAGATGGCAACTGGGATGCT 464

Qy      241 TGGGGCGACTGGAGTGACTGCTCCCGGACCTGTGGGGGAGGAGCATCATATTCTCTGCGG 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      465 TGGGGCGACTGGAGTGACTGCTCCCGGACCTGTGGGGGAGGAGCATCATATTCTCTGCGG 524

Qy      301 AGATGTTTGACTGGAAGGAATTGTGAAGGGCAGAACATTCGGTACAAGACATGCAGCAAT 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      525 AGATGTTTGACTGGAAGGAATTGTGAAGGGCAGAACATTCGGTACAAGACATGCAGCAAT 584

Qy      361 CATGACTGCCCTCCAGATGCAGAAGATTTCAGAGCCCAGCAGTGCTCAGCCTACAATGAT 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      585 CATGACTGCCCTCCAGATGCAGAAGATTTCAGAGCCCAGCAGTGCTCAGCCTACAATGAT 644

Qy      421 GTCCAGTATCAGGGGCATTACTATGAATGGCTTCCACGATATAATGATCCTGCTGCCCCG 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      645 GTCCAGTATCAGGGGCATTACTATGAATGGCTTCCACGATATAATGATCCTGCTGCCCCG 704

Qy      481 TGTGCACTCAAGTGTCATGCACAAGGACAAAACTTGGTGGTGGAGCTGGCACCTAAGGTA 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      705 TGTGCACTCAAGTGTCATGCACAAGGACAAAACTTGGTGGTGGAGCTGGCACCTAAGGTA 764

Qy      541 CTGGATGGAACTCGTTGCAACACGGACTCCTTGGACATGTGTATCAGTGGCATCTGTCAG 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      765 CTGGATGGAACTCGTTGCAACACGGACTCCTTGGACATGTGTATCAGTGGCATCTGTCAG 824

Qy      601 GCAGTGGGCTGCGATCGGCAACTGGGAAGCAATGCCAAGGAGGACAACTGTGGAGTCTGT 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      825 GCAGTGGGCTGCGATCGGCAACTGGGAAGCAATGCCAAGGAGGACAACTGTGGAGTCTGT 884

Qy      661 GCCGGCGATGGCTCCACCTGCAGGCTTGTACGGGGACAATCAAAGTCACACGTTTCTCCT 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      885 GCCGGCGATGGCTCCACCTGCAGGCTTGTACGGGGACAATCAAAGTCACACGTTTCTCCT 944

Qy      721 GAAAAAAGAGAAGAAAATGTAATTGCTGTTCCTTTGGGAAGTCGAAGTGTGAGAATTACA 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      945 GAAAAAAGAGAAGAAAATGTAATTGCTGTTCCTTTGGGAAGTCGAAGTGTGAGAATTACA 1004

Qy      781 GTGAAAGCACCTGCCCACCTCTTTATTGAATCAAAAACACTTCAACGAAGCAAAGGAGAA 840
            ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db     1005 GTGAAAGCACCTGTCCACCTCTTTATTGAATCAAAAACACTTCAACGAAGCAAAGGAGAA 1064

Qy      841 CACAGCTTTAACAGCCCCCGGGTCTTTGTCGTAGAAAACACAACAGTGGAATTTCAGAGG 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1065 CACAGCTTTAACAGCCCCCGGGTCTTTGTCGTAGAAAACACAACAGTGGAATTTCAGAGG 1124

Qy      901 GGCTCCGAGAGGCAAACTTTTAACATTCCAGGACCTCTGATGGCTGATTTCATCTTCAAG 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1125 GGCTCCGAGAGGCAAACTTTTAACATTCCAGGACCTCTGATGGCTGATTTCATCTTCAAG 1184

Qy      961 ACCAGGTACACTGCAGCCAAAGACAGCGTGGTTCAGTTCTTCTTTTACCAGCCCATCAGT 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1185 ACCAGGTACACTGCAGCCAAAGACAGGCTGGTTCAGTTCTTCTTTTACCAGCCCATCAGT 1244

Qy       1021 CATCAGTGGAGACAAACTGACTTCTTTCCCTGCACTGTGACGTGTGGAGGAGGTTATCAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1245 CATCAGTGGAGACAAACTGACTTCTTTCCCTGCACTGTGACGTGTGGAGGAGGTTATCAG 1304

Qy       1081 CTCAATTCTGCTGAATGTGTGGATATCCGCTTGAAGAGGGTAGTTCCTGACCATTATTGG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1305 CTCAATTCTGCTGAATGTGTGGATATCCGCTTGAAGAGGGTAGTTCCTGACCATTATTGG 1364

Qy       1141 CACTACTACCCTGAAAATGTAAAACCAAAACCAAAACTGAAGGAATGCAGCATGGATCCC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1365 CACTACTACCCTGAAAATGTAAAACCAAAACCAAAACTGAAGGAATGCAGCATGGATCCC 1424

Qy       1201 TGCCCATCAAGTGATGGATTTAAAGAGATAATGCCCTATGACCACTTCCAACCTCTCCCT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1425 TGCCCATCAAGTGATGGATTTAAAGAGATAATGCCCTATGACCACTTCCAACCTCTCCCT 1484

Qy       1261 CCCTGGGAACATAATCCTTGGACTGCATGTTCCGTGTCCTGTGGAGGAGGGATTCAGAGA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1485 CCCTGGGAACATAATCCTTGGACTGCATGTTCCGTGTCCTGTGGAGGAGGGATTCAGAGA 1544

Qy       1321 CGGAGCTTTGTGTGTGTAGAGGAATCCATCCATGGAGAGATATTGCAGGTGGAAGAATGG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1545 CGGAGCTTTGTGTGTGTAGAGGAATCCATCCATGGAGAGATATTGCAGGTGGAAGAATGG 1604

Qy       1381 AAGTGCATGTACGGACCCAAACCCAAGGTTATGCAAACTTGTAATCTGTTGGATTGCCCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1605 AAGTGCATGTACGGACCCAAACCCAAGGTTATGCAAACTTGTAATCTGTTGGATTGCCCC 1664

Qy       1441 AAGTGGATTGCCCATGGAGGGTTGCAGTGCACAGTGACTTGTGGCCGAGGGTTACGGTAC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1665 AAGTGGATTGCCCATGGAGGGTTGCAGTGCACAGTGACTTGTGGCCGAGGGTTACGGTAC 1724

Qy       1501 CGGGTTGTTCTGTGTATTAACTACCGGGAGAGCATGTTGGGGGCTGCAATCCACAACTG 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1725 CGGGTTGTTCTGTGTATTAACTACCGGGAGAGCATGTTGGGGGCTGCAATCCACAACTG 1784

Qy       1561 AAGTTACACATCAAAGAAGAATGTGTCATCCCATCCCGTGTTATAAACCAAAAGAAAAA 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1785 AAGTTACACATCAAAGAAGAATGTGTCATCCCATCCCGTGTTATAAACCAAAAGAAAAA 1844

Qy       1621 AGTCCAGTGGAAGCAAAATTGCCTTGGCTGAAACAAGCACAAGAACTAGAAGAGACCAGA 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1845 AGTCCAGTGGAAGCAAAATTGCCTTGGCTGAAACAAGCACAAGAACTAGAAGAGACCAGA 1904

Qy       1681 ATAGCAACAGAAGAACCAACGTTCATTCCAGAACCCTGGTCAGCCTGCAGTACCACGTGT 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1905 ATAGCAACAGAAGAACCAACGTTCATTCCAGAACCCTGGTCAGCCTGCAGTACCACGTGT 1964

Qy       1741 GGGCCAGGTGTGCAGGTCCGGGAGGTGAAGTGCGTGTGCTCCTCACATTCACGCAGACT 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1965 GGGCCAGGTGTGCAGGTCCGGGAGGTGAAGTGCGTGTGCTCCTCACATTCACGCAGACT 2024

Qy       1801 GAGACTGAGCTGCCCGAGGAAGAGTGTGAAGGCCCCAAGCTGCCCACCGAACGGCCCTGC 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2025 GAGACTGAGCTGCCCGAGGAAGAGTGTGAAGGCCCCAAGCTGCCCACCGAACGGCCCTGC 2084

Qy       1861 CTCCTGGAAGCATGTGATGAGAGCCCGGCCTCCCGAGAGCTAGACATCCCTCTCCCTGAG 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2085 CTCCTGGAAGCATGTGATGAGAGCCCGGCCTCCCGAGAGCTAGACATCCCTCTCCCTGAG 2144

Qy       1921 GACAGTGAGACGACTTACGACTGGGAGTACGCTGGGTTCACCCCTTGCACAGCAACATGC 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2145 GACAGTGAGACGACTTACGACTGGGAGTACGCTGGGTTCACCCCTTGCACAGCAACATGC 2204

Qy       1981 TTGGGAGGCCATCAAGAAGCCATAGCAGTGTGCTTACATATCCAGACCCAGCAGACAGTC 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2205 TTGGGAGGCCATCAAGAAGCCATAGCAGTGTGCTTACATATCCAGACCCAGCAGACAGTC 2264

Qy       2041 AATGACAGCTTGTGTGATATGGTCCACCGTCCTCCAGCCATGAGCCAGGCCTGTAACACA 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2265 AATGACAGCTTGTGTGATATGGTCCACCGTCCTCCAGCCATGAGCCAGGCCTGTAACACA 2324

Qy       2101 GAGCCCTGTCCCCCAGGTGGCATGTGGGCTCTTGGGGGGCCTGGCTACCTACCTGGTGGA 2160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2325 GAGCCCTGTCCCCCAGGTGGCATGTGGGCTCTTGGGGGGCCTGGCTACCTACCTGGTGGA 2384

Qy       2161 GTTGGAATTCAGACCCGAGATGTGTACTGCCTGCACCAGGGGAGACCCCTGCCCCTGCCT 2220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2385 GTTGGAATTCAGACCCGAGATGTGTACTGCCTGCACCAGGGGAGACCCCTGCCCCTGCCT 2444

Qy       2221 GAGGAGTGCCGAGATGAAAACCCCATGCTTTACAAGCATCAATCATTTGATGCCCT 2280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2445 GAGGAGTGCCGAGATGAAAACCCCATGCTTTACAAGCATCAATCATTTGATGCCCT 2504

Qy       2281 CCTGGCTGGCACATTGAAG 2299
              |||||||||||||||||||
Db       2505 CCTGGCTGGCACATTGAAG 2523


RESULT 15
AX224821
LOCUS       AX224821                2316 bp    DNA     linear   PAT 10-SEP-2001
DEFINITION  Sequence 13 from Patent WO0161011.
ACCESSION   AX224821
VERSION     AX224821.1  GI:15554915
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2316)
  AUTHORS   Donoho,G., Scoville,J., Turner,C.A., Friedrich,G., Zambrowicz,B.
            and Sands,A.T.
  TITLE     Novel human thrombospondin repeat proteins and polynucleotides
            encoding the same
  JOURNAL   Patent: WO 0161011-A 13 23-AUG-2001;
            Lexicon Genetics Incorporated (US)
FEATURES             Location/Qualifiers
     source          1..2316
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      617 a    575 c    616 g    508 t
ORIGIN

  Query Match             42.9%;  Score 2180;  DB 6;  Length 2316;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 2180;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        311 CTGGAAGGAATTGTGAAGGGCAGAACATTCGGTACAAGACATGCAGCAATCATGACTGCC 370
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         89 CTGGAAGGAATTGTGAAGGGCAGAACATTCGGTACAAGACATGCAGCAATCATGACTGCC 148

Qy        371 CTCCAGATGCAGAAGATTTCAGAGCCCAGCAGTGTTCAGCCTACAATGATGTCCAGTATC 430
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        149 CTCCAGATGCAGAAGATTTCAGAGCCCAGCAGTGTTCAGCCTACAATGATGTCCAGTATC 208

Qy        431 AGGGGCATTACTATGAATGGCTTCCACGATATAATGATCCTGCTGCCCCGTGTGCACTCA 490
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        209 AGGGGCATTACTATGAATGGCTTCCACGATATAATGATCCTGCTGCCCCGTGTGCACTCA 268

Qy        491 AGTGTCATGCACAAGGACAAAACTGGGTGGTGGAGCTGGCACCTAAGGTACTGGATGGAA 550
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        269 AGTGTCATGCACAAGGACAAAACTGGGTGGTGGAGCTGGCACCTAAGGTACTGGATGGAA 328

Qy        551 CTCGTTGCAACACGGACTCCTTGGACATGTGTATCAGTGGCATCTGTCAGGCAGTGGGCT 610
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        329 CTCGTTGCAACACGGACTCCTTGGACATGTGTATCAGTGGCATCTGTCAGGCAGTGGGCT 388

Qy        611 GCGATCGGCAACTGGGAAGCAATGCCAAGGAGGACAACTGTGGAGTCTGTGCCGGCGATG 670
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        389 GCGATCGGCAACTGGGAAGCAATGCCAAGGAGGACAACTGTGGAGTCTGTGCCGGCGATG 448

Qy        671 GCTCCACCTGCAGGCTTGTACGGGGACAATCAAACTCAACGGTTTCTCCTGAAAAAAGAG 730
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      449 GCTCCACCTGCAGGCTTGTACGGGGACAATCAAAGTCACACGTTTCTCCTGAAAAAAGAG 508

Qy      731 AAGAAAATGTAATTGCTGTTCCTTTGGGAAGTCGAAGTGTGAGAATTACAGTGAAAGGAC 790
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      509 AAGAAAATGTAATTGCTGTTCCTTTGGGAAGTCGAAGTGTGAGAATTACAGTGAAAGGAC 568

Qy      791 CTGCCCACCTCTTTATTGAATCAAAAACACTTCAAGGAAGCAAAGGAGAACACAGCTTTA 850
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      569 CTGCCCACCTCTTTATTGAATCAAAAACACTTCAAGGAAGCAAAGGAGAACACAGCTTTA 628

Qy      851 ACAGCCCCGGCGTCTTTGTCGTAGAAAACACAACAGTGGAATTTCAGAGGGGCTCCGAGA 910
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      629 ACAGCCCCGGCGTCTTTGTCGTAGAAAACACAACAGTGGAATTTCAGAGGGGCTCCGAGA 688

Qy      911 GGCAAACTTTTAAGATTCCAGGACCTCTGATGGCTGATTTCATCTTCAAGACCAGGTACA 970
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      689 GGCAAACTTTTAAGATTCCAGGACCTCTGATGGCTGATTTCATCTTCAAGACCAGGTACA 748

Qy      971 CTGCAGCCAAAGACAGCGTGGTTCAGTTCTTCTTTTACCAGCCCATCAGTCATCAGTGGA 1030
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      749 CTGCAGCCAAAGACAGCGTGGTTCAGTTCTTCTTTTACCAGCCCATCAGTCATCAGTGGA 808

Qy     1031 GACAAACTGACTTCTTTCCCTGCACTGTGACGTGTGGAGGAGGTTATCAGCTCAATTCTG 1090
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      809 GACAAACTGACTTCTTTCCCTGCACTGTGACGTGTGGAGGAGGTTATCAGCTCAATTCTG 868

Qy     1091 CTGAATGTGTGGATATCCGCTTGAAGAGGGTAGTTCCTGACCATTATTGTCACTACTACC 1150
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      869 CTGAATGTGTGGATATCCGCTTGAAGAGGGTAGTTCCTGACCATTATTGTCACTACTACC 928

Qy     1151 CTGAAAATGTAAAACCAAAACCAAAACTGAAGGAATGCAGCATGGATCCCTGCCCATCAA 1210
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      929 CTGAAAATGTAAAACCAAAACCAAAACTGAAGGAATGCAGCATGGATCCCTGCCCATCAA 988

Qy     1211 GTGATGGATTTAAAGAGATAATGCCCTATGACCACTTCCAACCTCTTCCTCGCTGGGAAC 1270
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      989 GTGATGGATTTAAAGAGATAATGCCCTATGACCACTTCCAACCTCTTCCTCGCTGGGAAC 1048

Qy     1271 ATAATCCTTGGACTGCATGTTCCGTGTCCTGTGGAGGAGGGATTCAGAGACGGAGCTTTG 1330
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1049 ATAATCCTTGGACTGCATGTTCCGTGTCCTGTGGAGGAGGGATTCAGAGACGGAGCTTTG 1108

Qy     1331 TGTGTGTAGAGGAATCCATGCATGGAGAGATATTGCAGGTGGAAGAATGGAAGTGCATGT 1390
            |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db     1109 TGTGTGTAGAGGAATCCATGCATGGAGAGATATTGCAGGTGGAAGAATGGAAGTGCATGT 1168

Qy     1391 ACGCACCCAAACCCAAGGTTATGCAAACTTGTAATCTGTTTGATTGCCCCAAGTGGATTG 1450
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1169 ACGCACCCAAACCCAAGGTTATGCAAACTTGTAATCTGTTTGATTGCCCCAAGTGGATTG 1228

Qy     1451 CCATGGAGTGGTCTCAGTGCACAGTGACTTGTGGCCGAGGGTTACGGTACCGGGTTGTTC 1510
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1229 CCATGGAGTGGTCTCAGTGCACAGTGACTTGTGGCCGAGGGTTACGGTACCGGGTTGTTC 1288

Qy     1511 TGTGTATTAACCACCGCGGAGAGCATGTTGGGGGCTGCAATCCACAACTGAAGTTACACA 1570
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1289 TGTGTATTAACCACCGCGGAGAGCATGTTGGGGGCTGCAATCCACAACTGAAGTTACACA 1348

Qy     1571 TCAAAGAAGAATGTGTCATTCCCATCCCGTGTTATAAACCAAAAGAAAAAAGTCCAGTGG 1630
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1349 TCAAAGAAGAATGTGTCATTCCCATCCCGTGTTATAAACCAAAAGAAAAAAGTCCAGTGG 1408

Qy     1631 AAGCAAAATTGCCTTGGCTGAAACAAGCACAAGAACTAGAAGAGACCAGAATAGCAACAG 1690
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1409 AAGCAAAATTGCCTTGGCTGAAACAAGCACAAGAACTAGAAGAGACCAGAATAGCAACAG 1468

Qy     1691 AAGAACCAACGTTCATTCCAGAACCCTGGTCAGCCTGCAGTACCACGTGTGGGCCAGGTG 1750
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1469 AAGAACCAACGTTCATTCCAGAACCCTGGTCAGCCTGCAGTACCACGTGTGGGCCAGGTG 1528

Qy     1751 TGCAGGTCCGCGAGGTGAAGTGCCGTGTGCTCCTCACATTCACGCAGACTGAGACTGAGC 1810
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1529 TGCAGGTCCGCGAGGTGAAGTGCCGTGTGCTCCTCACATTCACGCAGACTGAGACTGAGC 1588

Qy     1811 TGCCCGAGGAAGAGTGTGAAGGCCCCAAGCTGCCCACCGAACGGCCCTGCCTCCTGGAAG 1870
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1589 TGCCCGAGGAAGAGTGTGAAGGCCCCAAGCTGCCCACCGAACGGCCCTGCCTCCTGGAAG 1648

Qy     1871 CATGTGATGAGAGCCCGGCCTCCCGAGAGCTAGACATCCCTCTCCCTGAGGACAGTGAGA 1930
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1649 CATGTGATGAGAGCCCGGCCTCCCGAGAGCTAGACATCCCTCTCCCTGAGGACAGTGAGA 1708

Qy     1931 CGACTTACGACTGGGAGTACGCTGGGTTCACCCCTTGCACAGCAACATGCTTGGGAGGCC 1990
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1709 CGACTTACGACTGGGAGTACGCTGGGTTCACCCCTTGCACAGCAACATGCTTGGGAGGCC 1768

Qy     1991 ATCAAGAAGCCATAGCAGTGTGCTTACATATCCAGACCCAGCAGACAGTCAATGACAGCT 2050
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1769 ATCAAGAAGCCATAGCAGTGTGCTTACATATCCAGACCCAGCAGACAGTCAATGACAGCT 1828

Qy     2051 TGTGTGATATGGTCCACCGTCCTCCAGCCATGAGCCAGGCCTGTAACACAGAGCCCTGTC 2110
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1829 TGTGTGATATGGTCCACCGTCCTCCAGCCATGAGCCAGGCCTGTAACACAGAGCCCTGTC 1888

Qy     2111 CCCCCAGGTGGCATGTGGGCTCTTGGGGGCCCTGCTCAGCTACCTGTGGAGTTGGAATTC 2170
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1889 CCCCCAGGTGGCATGTGGGCTCTTGGGGGCCCTGCTCAGCTACCTGTGGAGTTGGAATTC 1948

Qy     2171 AGACCCGAGATGTGTACTGCCTGCACCCAGGGGAGACCCCTGCCCCTCCTGAGGAGTGCC 2230
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1949 AGACCCGAGATGTGTACTGCCTGCACCCAGGGGAGACCCCTGCCCCTCCTGAGGAGTGCC 2008

Qy     2231 GAGATGAAAAGCCCCATGCTTTACAAGCATGCAATCAGTTTGACTGCCCTCCTGGCTGGC 2290
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2009 GAGATGAAAAGCCCCATGCTTTACAAGCATGCAATCAGTTTGACTGCCCTCCTGGCTGGC 2068

Qy     2291 ACATTGAAGAATGGCAGCAGTGTTCCAGGACTTGTGGCGGGGGAACTCAGAACAGAAGAG 2350
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2069 ACATTGAAGAATGGCAGCAGTGTTCCAGGACTTGTGGCGGGGGAACTCAGAACAGAAGAG 2128

Qy     2351 TCACCTGTCGGCAGCTGCTAACGGATGGCAGCTTTTTGAATCTCTCAGATGAATTGTGCC 2410
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2129 TCACCTGTCGGCAGCTGCTAACGGATGGCAGCTTTTTGAATCTCTCAGATGAATTGTGCC 2188

Qy     2411 AAGGACCCAAGGCATCGTCTCACAAGTCCTGTGCCAGGACAGACTGTCCTCCACATTTAG 2470
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2189 AAGGACCCAAGGCATCGTCTCACAAGTCCTGTGCCAGGACAGACTGTCCTCCACATTTAG 2248

Qy     2471 CTGTGGGAGACTGGTCGAAG 2490
            ||||||||||||||||||||
Db     2249 CTGTGGGAGACTGGTCGAAG 2268


Search completed: March 22, 2003, 09:31:57
Job time : 9924 secs
```

```
                    GenCore version 5.1.4_p5_4578
              Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:          March 22, 2003, 02:26:33 ; Search time 776 Seconds
                                            (without alignments)
                                            14730.855 Million cell updates/sec

Title:           US-09-784-358-1
Perfect score:   5076
Sequence:        1 atgggcttcctagcagcccc.........aagtcatatgtcaagagatatag 5076

Scoring table:   IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0

Searched:        2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters:      4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       N_Geneseq_101002:*
                 1:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
                 2:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
                 3:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
                 4:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
                 5:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
                 6:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
                 7:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
                 8:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
                 9:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
                 10:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
                 11:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
                 12:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
                 13:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
                 14:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
                 15:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
                 16:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
                 17:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
                 18:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
                 19:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
                 20:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
                 21:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
                 22:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
                 23:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
                 24:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                SUMMARIES
                         %
Result                 Query
  No.     Score     Match Length DB  ID                       Description
                                                                                         
    1      5076     100.0   5076  22  AAD14364                 Novel human protei
    2    5069.6      99.9   7260  22  AAD17757                 Human novel KIAA12
    3      4766      93.9   4854  22  AAD14371                 Novel human protei
    4    3862.2      76.1   6303  22  AAD17758                 Human novel KIAA12
    5      3859      76.0   6294  22  AAD17756                 Human novel KIAA12
    6      3755      74.0   8578  22  AAD14372                 Novel human protei
    7    3289.6      64.8   5343  21  AAC74509                 Human ORFX ORF264
    8    2862.8      56.4   2912  23  AAF87125                 NOV14 coding seque
    9      2490      49.1   2538  22  AAD14369                 Novel human protei
   10      2180      42.9   2316  22  AAD14370                 Novel human protei
   11      2119      41.7   2175  22  AAD14367                 Novel human protei
   12      1809      35.6   1953  22  AAD14368                 Novel human protei
   13      1327      26.1   1341  22  AAD14365                 Novel human protei
   14      1017      20.0   1119  22  AAD14366                 Novel human protei
   15     925.4      18.2   5769  22  ABA06408                 Human cDNA SEQ II
   16     924.2      18.2   5720  22  AAF97896                 Human secreted pro
   17       913      18.0   2689  21  AAA47458                 Human TANGO 224 co
   18       913      18.0   2762  23  ABV20967                 Human prostate exp
   19       913      18.0   2762  23  ABV26815                 Human prostate exp
   20       912      18.0   2689  21  AAA47482                 Human TANGO 224 co
   21       846      16.7    846  22  ABA08559                 Human HSPP 124 pro
   22     843.2      16.6   3559  22  AAF97926                 Human secreted pro
   23       828      16.3   3045  21  AAZ98232                 Human signal pepti
   24       738      14.5   1869  20  AAZ34146                 Human PRO1071 nucl
   25       738      14.5   1869  21  AAC78532                 Human PRO1071 (UNQ
   26       738      14.5   1869  22  AAS46142                 Human DNA encoding
   27       738      14.5   1869  24  ABL95632                 Human angiogenesis
   28       738      14.5   1869  24  ABL88143                 Human PRO1071 cDNA
   29       738      14.5   1869  24  ABK33612                 cDNA encoding huma
   30     733.2      14.4   1863  22  AAF63447                 Human ADAMTS-E1 cD
   31     584.8      11.5   1299  21  AAC75770                 Human ORFX ORF1527
   32       536      10.6   1616  22  AAI66429                 Human thrombotic p
   33     415.8       8.2    996  20  AAX97684                 Extended human sec
   34     352.2       6.9    852  22  AAF97927                 Human secreted pro
   35     351.6       6.9    747  20  AAX97583                 Extended human sec
   36     232.4       4.6    861  22  AAH99417                 Human protein enco
   37     202.2       4.0    405  22  ABA08536                 Human HSPP 124 pro
   38     177.6       3.5   1107  22  ABA08462                 Human BPTI/Kunitz
   39     112.4       2.2   2274  22  AAD07967                 Human protease rel
   40     112.4       2.2   2727  22  AAD07956                 Human protease rel
   41     112.4       2.2   2879  22  AAH49370                 Human metalloprote
   42     109.8       2.2   2848  22  AAF63439                 Human ADAMTS-6 cDN
   43     106.6       2.1   3169  22  AAD07968                 Human protease rel
   44     104.2       2.1   2145  23  ABA94917                 Human metalloprote
   45     104.6       2.1   3133  24  ABA94919                 Human metalloprote

                                ALIGNMENTS

RESULT 1
AAD14364
ID   AAD14364 standard; cDNA; 5076 BP.
XX
AC   AAD14364;
XX
DT   01-NOV-2001  (first entry)
XX
DE   Novel human protein (NHP) cDNA #1.
XX
KW   Novel human protein; NHP; thrombospondin; gene therapy; cancer;
KW   cytostatic; antisense therapy; angiogenesis; biological disorder; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..5076
FT                   /*tag=  a
FT                   /product=  "Novel human protein"
XX
PN   WO200161011-A2.
XX
PD   23-AUG-2001.
XX
PF   15-FEB-2001; 2001WO-US05290.
XX
PR   17-FEB-2000; 2000US-0183282.
XX
PA   (LEXI-) LEXICON GENETICS INC.
XX
PI   Donoho G,  Scoville J,  Turner CA,  Friedrich G,  Zambrowicz B;
PI   Sands AT;
```

```
XX
DR   WPI; 2001-514776/56.
DR   P-PSDB; AAE07863.
XX
PT   An isolated nucleic acid encoding a thrombospondin useful as a
PT   hybridization probe and gene therapy treatments of cancer -
XX
PS   Claim 1; Page 29-30; 56pp; English.
XX
CC   The invention relates to novel human polynucleotides encoding proteins
CC   that share sequence similarity with animal proteins having thrombospondin
CC   repeats. NHPs are expressed in human cell lines, pituitary, lymph nodes,
CC   prostrate, testis, adrenal gland, uterus, foetal kidney, foetal lung and
CC   gene trapped human cells. NHPs are thrombcspondins useful for treating
CC   biological disorders involving angiogenesis, cancer and development and
CC   also in pharmacogenomic applications. NHPs are useful as a hybridisation
CC   probe for screening libraries, assessing gene expression patterns and
CC   also in gene therapy. Proteins having thrombospondin repeats act as
CC   receptors, secreted extracellular matrix proteins and proteases. The
CC   present sequence is a cDNA encoding novel human protein (NHP).
XX
SQ   Sequence 5076 BP; 1384 A; 1227 C; 1338 G; 1127 T; 0 other;

  Query Match             100.0%;  Score 5076;  DB 22;  Length 5076;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 5076;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGGCTTCCTGGACGAGCCCCTGGTGGGTGCTGATAGGGATGGTCTTCATGCACTCTCCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGCTTCCTGGACGAGCCCCTGGTGGGTGCTGATAGGGATGGTCTTCATGCACTCTCCC 60

Qy         61 CTCCCGCAGACCACAGCTGAGAAATCTCCTGGAGCCTATTTCCTTCCCGAGTTTGCACTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTCCCGCAGACCACAGCTGAGAAATCTCCTGGAGCCTATTTCCTTCCCGAGTTTGCACTT 120

Qy        121 TCTCCTCAGGGAAGTTTTCTGGAAGACACAACAGGGGAGCAGTTCCTCACTTATCGCTAT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCTCCTCAGGGAAGTTTTCTGGAAGACACAACAGGGGAGCAGTTCCTCACTTATCGCTAT 180

Qy        181 GATGACCAGACCTCAAGAAACACTCGTTCAGATGAAGACAAAGATGGCAACTGGGATGCT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GATGACCAGACCTCAAGAAACACTCGTTCAGATGAAGACAAAGATGGCAACTGGGATGCT 240

Qy        241 TGGGGCGACTGGAGTGACTGCTCCCGGACCTGTGGGGGAGGAGCATCATATTCTCTGCGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGGGGCGACTGGAGTGACTGCTCCCGGACCTGTGGGGGAGGAGCATCATATTCTCTGCGG 300

Qy        301 AGATGTTTGACTGGAAGGAATTGTGAAGGGCAGAACATTCGGTACAAGACATGCAGCAAT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGATGTTTGACTGGAAGGAATTGTGAAGGGCAGAACATTCGGTACAAGACATGCAGCAAT 360

Qy        361 CATGACTGCCCTCCAGATGCAGAAGATTTCAGAGCCCAGCAGTGCTCAGCCTACAATGAT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CATGACTGCCCTCCAGATGCAGAAGATTTCAGAGCCCAGCAGTGCTCAGCCTACAATGAT 420

Qy        421 GTCCAGTATCAGGGGCATTACTATGAATGGCTTCCACGATATAATGATCCTGCTGCCCCG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GTCCAGTATCAGGGGCATTACTATGAATGGCTTCCACGATATAATGATCCTGCTGCCCCG 480

Qy        481 TGTGCACTCAAGTGTCATGCACAAGGACAAAACTTGGTGGTGGAGCTGGCACCTAAGGTA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TGTGCACTCAAGTGTCATGCACAAGGACAAAACTTGGTGGTGGAGCTGGCACCTAAGGTA 540

Qy        541 CTGGATGGAACTCGTTGCAACACGGACTCCTTGGACATGTGTATCAGTGGCATCTGTCAG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTGGATGGAACTCGTTGCAACACGGACTCCTTGGACATGTGTATCAGTGGCATCTGTCAG 600

Qy        601 GCAGTGGGCTGCGATCGGCAACTGGGAAGCAATGCCAAGGAGGACAACTGTGGAGTCTGT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GCAGTGGGCTGCGATCGGCAACTGGGAAGCAATGCCAAGGAGGACAACTGTGGAGTCTGT 660

Qy        661 GCCGGCGATGGCTCCACCTGCAGGCTTGTACGGGGACAATCAAAGTCACACGTTTCTCCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GCCGGCGATGGCTCCACCTGCAGGCTTGTACGGGGACAATCAAAGTCACACGTTTCTCCT 720

Qy        721 GAAAAAAGAGAAGAAAATGTAATTGCTGTTCCTTTGGGAAGTCGAAGTGTGAGAATTACA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GAAAAAAGAGAAGAAAATGTAATTGCTGTTCCTTTGGGAAGTCGAAGTGTGAGAATTACA 780

Qy        781 GTGAAAGGACCTGCCCACCTCTTTATTGAATCAAAAACACTTCAAGGAAGCAAAGGAGAA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GTGAAAGGACCTGCCCACCTCTTTATTGAATCAAAAACACTTCAAGGAAGCAAAGGAGAA 840

Qy        841 CACAGCTTTAACAGCCCCGGCGTCTTTGTCGTAGAAAACACAACAGTGGAATTTCAGAGG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CACAGCTTTAACAGCCCCGGCGTCTTTGTCGTAGAAAACACAACAGTGGAATTTCAGAGG 900

Qy        901 GGCTCCGAGAGGCAAACTTTTAAGATTCCAGGACCTCTGATGGCTGATTTCATCTTCAAG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GGCTCCGAGAGGCAAACTTTTAAGATTCCAGGACCTCTGATGGCTGATTTCATCTTCAAG 960

Qy        961 ACCAGGTACACTGCAGCCAAAGACAGCGTGGTTCAGTTCTTCTTTTACCAGCCCATCAGT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 ACCAGGTACACTGCAGCCAAAGACAGCGTGGTTCAGTTCTTCTTTTACCAGCCCATCAGT 1020

Qy       1021 CATCAGTGGAGACAAACTGACTTCTTTCCCTGCACTGTGACGTGTGGAGGAGGTTATCAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CATCAGTGGAGACAAACTGACTTCTTTCCCTGCACTGTGACGTGTGGAGGAGGTTATCAG 1080

Qy       1081 CTCAATTCTGCTGAATGTGTGGATATCCGCTTGAAGAGGGTAGTTCCTGACCATTATTGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CTCAATTCTGCTGAATGTGTGGATATCCGCTTGAAGAGGGTAGTTCCTGACCATTATTGT 1140

Qy       1141 CACTACTACCCTGAAAATGTAAAACCAAAACCAAAACTGAAGGAATGCAGCATGGATCCC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CACTACTACCCTGAAAATGTAAAACCAAAACCAAAACTGAAGGAATGCAGCATGGATCCC 1200

Qy       1201 TGCCCATCAAGTGATGGATTTAAAGAGATAATGCCCTATGACCACTTCCAACCTCTTCCT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TGCCCATCAAGTGATGGATTTAAAGAGATAATGCCCTATGACCACTTCCAACCTCTTCCT 1260

Qy       1261 CGCTGGGAACATAATCCTTGGACTGCATGTTCCGTGTCCTGTGGAGGAGGGATTCAGAGA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CGCTGGGAACATAATCCTTGGACTGCATGTTCCGTGTCCTGTGGAGGAGGGATTCAGAGA 1320

Qy       1321 CGGAGCTTTGTGTGTGTAGAGGAATCCATGCATGGAGAGATATTGCAGGTGGAAGAATGG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CGGAGCTTTGTGTGTGTAGAGGAATCCATGCATGGAGAGATATTGCAGGTGGAAGAATGG 1380

Qy       1381 AAGTGCATGTACGCACCCAAACCCAAGGTTATGCAAACTTGTAATCTGTTTGATTGCCCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 AAGTGCATGTACGCACCCAAACCCAAGGTTATGCAAACTTGTAATCTGTTTGATTGCCCC 1440

Qy       1441 AAGTGGATTGCCATGGAGTGGTCTCAGTGCACAGTGACTTGTGGCCGAGGGTTACGGTAC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 AAGTGGATTGCCATGGAGTGGTCTCAGTGCACAGTGACTTGTGGCCGAGGGTTACGGTAC 1500

Qy       1501 CGGGTTGTTCTGTGTATTAACCACCGCGGAGAGCATGTTGGGGGCTGCAATCCACAACTG 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 CGGGTTGTTCTGTGTATTAACCACCGCGGAGAGCATGTTGGGGGCTGCAATCCACAACTG 1560

Qy       1561 AAGTTACACATCAAAGAAGAATGTGTCATTCCCATCCCGTGTTATAAACCAAAAGAAAAA 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 AAGTTACACATCAAAGAAGAATGTGTCATTCCCATCCCGTGTTATAAACCAAAAGAAAAA 1620

Qy       1621 AGTCCAGTGGAAGCAAAATTGCCTTGGCTGAAACAAGCACAAGAACTAGAAGAGACCAGA 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 AGTCCAGTGGAAGCAAAATTGCCTTGGCTGAAACAAGCACAAGAACTAGAAGAGACCAGA 1680

Qy       1681 ATAGCAACAGAAGAACCAACGTTCATTCCAGAACCCTGGTCAGCCTGCAGTACCACGTGT 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 ATAGCAACAGAAGAACCAACGTTCATTCCAGAACCCTGGTCAGCCTGCAGTACCACGTGT 1740

Qy       1741 GGGCCAGGTGTGCAGGTCCGCGAGGTGAAGTGCCGTGTGCTCCTCACATTCACGCAGACT 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 GGGCCAGGTGTGCAGGTCCGCGAGGTGAAGTGCCGTGTGCTCCTCACATTCACGCAGACT 1800
```

```
Qy       1801 GAGACTGAGCTGCCCGAGGAAGAGTGTGAAGGCCCCAAGCTGCCACCGAACGGCCCCTGC 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 GAGACTGAGCTGCCCGAGGAAGAGTGTGAAGGCCCCAAGCTGCCACCGAACGGCCCCTGC 1860

Qy       1861 CTCCTGGAAGCAGCGGATGAGAGGCCGGCCTCCCGAGAGCTAGACATCCCTCTCCCTGAG 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 CTCCTGGAAGCAGCGGATGAGAGGCCGGCCTCCCGAGAGCTAGACATCCCTCTCCCTGAG 1920

Qy       1921 GACAGTGAGACGACTTACGACTGGGAGTACGCTGGGTTCACCCCTTGCACAGCAACATGC 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 GACAGTGAGACGACTTACGACTGGGAGTACGCTGGGTTCACCCCTTGCACAGCAACATGC 1980

Qy       1981 TTGGGAGGCCATCAAGAAGCATAGCAGTGTGCTTACATATCCAGACCCAGCAGACAGTCC 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 TTGGGAGGCCATCAAGAAGCATAGCAGTGTGCTTACATATCCAGACCCAGCAGACAGTCC 2040

Qy       2041 AATGACAGCTTGTGTGGATATGGTCACCGTCCTCCAGCCATGAGCCAGGCCTGTAACACA 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 AATGACAGCTTGTGTGGATATGGTCACCGTCCTCCAGCCATGAGCCAGGCCTGTAACACA 2100

Qy       2101 GAGCCCTGTCCCCCCAGGTGGCATGTGGGCTCTTGGGGGCCCTGCTCAGCTACCTGTGGA 2160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 GAGCCCTGTCCCCCCAGGTGGCATGTGGGCTCTTGGGGGCCCTGCTCAGCTACCTGTGGA 2160

Qy       2161 GTTGGAATTCAGACCCGAGATGTGTACTGCCTGCACCCAGGGGAGACCCCTGCCCCTCCT 2220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 GTTGGAATTCAGACCCGAGATGTGTACTGCCTGCACCCAGGGGAGACCCCTGCCCCTCCT 2220

Qy       2221 GAGGACTGCCGAGATGAAAAGCCCCATGCTTTACAAGCATGCAATCAGTTTGACTGCCCT 2280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 GAGGACTGCCGAGATGAAAAGCCCCATGCTTTACAAGCATGCAATCAGTTTGACTGCCCT 2280

Qy       2281 CCTGGCTGGCACATTGAAGAATGGCAGCAGTGTTCCAGGACTTGTGGCCGGGGAACTCAG 2340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 CCTGGCTGGCACATTGAAGAATGGCAGCAGTGTTCCAGGACTTGTGGCCGGGGAACTCAG 2340

Qy       2341 AACAGAAGAGTCACCTGTCGGCAGCTGCTAACGGATGGCAGCTTTTTGAATCTCTCAGAT 2400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 AACAGAAGAGTCACCTGTCGGCAGCTGCTAACGGATGGCAGCTTTTTGAATCTCTCAGAT 2400

Qy       2401 GAATTGTGCCAAGGACCCAAGGCATCGTCTCACAAGTCCTGTGCCAGGACAGACTGTCCT 2460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 GAATTGTGCCAAGGACCCAAGGCATCGTCTCACAAGTCCTGTGCCAGGACAGACTGTCCT 2460

Qy       2461 CCACATTTAGCTGTGGGAGACTGGTCGAAGTGTTCTGTCAGTTGTGGTGTTGGAATCCAG 2520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 CCACATTTAGCTGTGGGAGACTGGTCGAAGTGTTCTGTCAGTTGTGGTGTTGGAATCCAG 2520

Qy       2521 AGAAGAAAGCAGGTGTTGTCAAAGCCTGGAGGCAAAGGTCGGGCATCCCCCTCAGTGAG 2580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 AGAAGAAAGCAGGTGTTGTCAAAGCCTGGAGGCAAAGGTCGGGCATCCCCCTCAGTGAG 2580

Qy       2581 ATGATGTGCAGGGATCTACCAGGGTTCCCTCTTGTAAGATCTTGCCAGATGCCTGAGTGC 2640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 ATGATGTGCAGGGATCTACCAGGGTTCCCTCTTGTAAGATCTTGCCAGATGCCTGAGTGC 2640

Qy       2641 AGTAAAATCAAATCAGAGATGAAGACAAAACTTGGTGAGCAGGGTCCGCAGATCCCTCAG 2700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 AGTAAAATCAAATCAGAGATGAAGACAAAACTTGGTGAGCAGGGTCCGCAGATCCCTCAG 2700

Qy       2701 GTCCAGAGAGTCTACATTCAGACAAGGGAAGAGAAGCGTATTAACCTGACCATTGGTAGC 2760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 GTCCAGAGAGTCTACATTCAGACAAGGGAAGAGAAGCGTATTAACCTGACCATTGGTAGC 2760

Qy       2761 AGAGCCTATTTGCTGCCCAACACATCCGTGATTATTAAGTGCCCCGTGCGACGATTCCAG 2820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 AGAGCCTATTTGCTGCCCAACACATCCGTGATTATTAAGTGCCCCGTGCGACGATTCCAG 2820

Qy       2821 AAATCTCTGATCCAGTGGGAGAAGGATGGCCGTTGCCTGCAGAACTCCAAACGGCTTGGC 2880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 AAATCTCTGATCCAGTGGGAGAAGGATGGCCGTTGCCTGCAGAACTCCAAACGGCTTGGC 2880

Qy       2881 ATCACCAAGTCAGGCTCACTAAAAATCCACGGCCTTGCTGCCCCCGACATCGGCGTGTAC 2940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 ATCACCAAGTCAGGCTCACTAAAAATCCACGGCCTTGCTGCCCCCGACATCGGCGTGTAC 2940

Qy       2941 CGGTGCATTGCAGGCTCTGCACAGGAAACAGTTGTGCTCAAGCTCATTGGTACTGACAAC 3000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 CGGTGCATTGCAGGCTCTGCACAGGAAACAGTTGTGCTCAAGCTCATTGGTACTGACAAC 3000

Qy       3001 CGGCTCATCGCACGCCCAGCCCTCAGGGAGCCTATGAGGGAATATCCTGGGATGGACCAC 3060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 CGGCTCATCGCACGCCCAGCCCTCAGGGAGCCTATGAGGGAATATCCTGGGATGGACCAC 3060

Qy       3061 AGCGAAGCCAATAGCTTGGGAGTCACATGGCACAAAATGAGGCAAATGTGGAATAACAAA 3120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 AGCGAAGCCAATAGCTTGGGAGTCACATGGCACAAAATGAGGCAAATGTGGAATAACAAA 3120

Qy       3121 AATGACCTTTATCTGGATGATGACCACATTAGTAACCAGCCTTTCTTGAGAGCTCTGTTA 3180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3121 AATGACCTTTATCTGGATGATGACCACATTAGTAACCAGCCTTTCTTGAGAGCTCTGTTA 3180

Qy       3181 GGCCACTGCAGCAATTCTGCAGGAAGCACCAACTCCTGGGAGTTGAAGAATAAGCAGTTT 3240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3181 GGCCACTGCAGCAATTCTGCAGGAAGCACCAACTCCTGGGAGTTGAAGAATAAGCAGTTT 3240

Qy       3241 GAAGCAGCAGTTAAACAAGGAGCATATAGCATGGATACAGCCCAGTTTGATGAGCTGATA 3300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3241 GAAGCAGCAGTTAAACAAGGAGCATATAGCATGGATACAGCCCAGTTTGATGAGCTGATA 3300

Qy       3301 AGAAACATGAGTCAGCTCATGGAAACCGGAGAGGTCAGCGATGATCTTGCGTCCCAGCTG 3360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3301 AGAAACATGAGTCAGCTCATGGAAACCGGAGAGGTCAGCGATGATCTTGCGTCCCAGCTG 3360

Qy       3361 ATATATCAGCTGGTGGCCGAATTAGCCAAGGCACAGCCAACACACATGCAGTGGGCGGGC 3420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3361 ATATATCAGCTGGTGGCCGAATTAGCCAAGGCACAGCCAACACACATGCAGTGGGCGGGC 3420

Qy       3421 ATCCAGGAAGAGACACCTCCTGCTGCTCAGCTCAGAGGGAAACAGGGAGTGTGTCCCCAA 3480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3421 ATCCAGGAAGAGACACCTCCTGCTGCTCAGCTCAGAGGGAAACAGGGAGTGTGTCCCCAA 3480

Qy       3481 AGCTCGGCATGCAAAAAACTCAGGCAAGCTGACATTCAAGCCGAAAGGACCTGTTCTCATG 3540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3481 AGCTCGGCATGCAAAAAACTCAGGCAAGCTGACATTCAAGCCGAAAGGACCTGTTCTCATG 3540

Qy       3541 AGGCAAAGCCAACCTCCCTCAATTTCATTTAATAAAACAATAAATTCCAGGATTGGAAAT 3600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3541 AGGCAAAGCCAACCTCCCTCAATTTCATTTAATAAAACAATAAATTCCAGGATTGGAAAT 3600

Qy       3601 ACAGTATACATTACAAAAAGGACAGAGGTCATCAATATACTGTGGGACCTTATCACCCCC 3660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3601 ACAGTATACATTACAAAAAGGACAGAGGTCATCAATATACTGTGGGACCTTATCACCCCC 3660

Qy       3661 AGTGAGGCCACATATACATGGACCAAGGATGGAACCTTGTTACAGCCCTCAGTAAAAATA 3720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3661 AGTGAGGCCACATATACATGGACCAAGGATGGAACCTTGTTACAGCCCTCAGTAAAAATA 3720

Qy       3721 ATTTTGGATGGAACTGGGAAGATACAGATACAGAATCCTACAAGGAAAGAACAAGGCATA 3780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3721 ATTTTGGATGGAACTGGGAAGATACAGATACAGAATCCTACAAGGAAAGAACAAGGCATA 3780

Qy       3781 TATGAATGTTCTGTAGCTAATCATCTTGGTTCAGATGTGGAAAGTTCTTCTGTGCTTGTAT 3840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3781 TATGAATGTTCTGTAGCTAATCATCTTGGTTCAGATGTGGAAAGTTCTTCTGTGCTTGTAT 3840

Qy       3841 GCAGAGGCACTGTCATCTTGTTGTGTTGAAAGAAATATCACCAAACCAGAGCACAACCAT 3900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3841 GCAGAGGCACTGTCATCTTGTTGTGTTGAAAGAAATATCACCAAACCAGAGCACAACCAT 3900

Qy       3901 CTGTCTGTTGTGGTTGGAGGCATCGTGGAGGCAGCCTTGGAGCAAACGTGACAATCCGA 3960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3901 CTGTCTGTTGTGGTTGGAGGCATCGTGGAGGCAGCCTTGGAGCAAACGTGACAATCCGA 3960

Qy       3961 GGTCCTGTAAAAGGTGTCCCCTCAGCCTAATATAACTTGGTTGAACAGAGGAGGACCTCTG 4020
```

```
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3961 TGTCCTGTAAAAGGTGTCCCTCAGCCTAATATAACTTGGTTGAAGAGAGGAGGATCTCTG 4020

Qy   4021 AGTGGCAATGTTTCCTTGCTTTTCAATGGATCCCTGTTGTTGCAGAATGTTTCCCTTGAA 4080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4021 AGTGGCAATGTTTCCTTGCTTTTCAATGGATCCCTGTTGTTGCAGAATGTTTCCCTTGAA 4080

Qy   4081 AATGAAGGAACCTACGTCTGCATAGCCACCAATGCTCTTGGAAAGGCAGTGGCAACATCT 4140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4081 AATGAAGGAACCTACGTCTGCATAGCCACCAATGCTCTTGGAAAGGCAGTGGCAACATCT 4140

Qy   4141 GTACTCCACTTGCTGGAACCAAGATGCCCAGACACTACAATCCTATTTCTCCAAGGACAT 4200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4141 GTACTCCACTTGCTGGAACCAAGATGCCCAGACACTACAATCCTATTTCTCCAAGGACAT 4200

Qy   4201 AAAAAGTACATTCTCCAGGCAACCAACACTAGAAGCAACAGCAATCACCCAACAGCACAA 4260
          ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db   4201 AAAAAGTACATTCTCCAGGCAACCAACACTAGAACCAACAGCAATCACCCAACAGCACAA 4260

Qy   4261 CCCCGGCCTCAAGAGCCTTTTTGGGAGCCTGGTAACTGGTCACATTGTTCTGCCACCTGT 4320
          ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db   4261 CCCCGGCCTCAAGAGCCTTTTTGGGAGCCTGGTAACTGGTCACATTGTTCTGCCACCTGT 4320

Qy   4321 GGTCATTTGGGAGCCCGCATTCAGAGACCCCAGTGTGTGATGGCCAATGGGCAGGAAGTG 4380
          ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db   4321 GGTCATTTGGGAGCCCGCATTCAGAGACCCCAGTGTGTGATGGCCAATGGGCAGGAAGTG 4380

Qy   4381 AGTGAGGCCCTGTGTGATCACCTCCAGAAGCCACTGGCTGGGTTTGAGCCCTGTAACATC 4440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4381 AGTGAGGCCCTGTGTGATCACCTCCAGAAGCCACTGGCTGGGTTTGAGCCCTGTAACATC 4440

Qy   4441 CGGGACTGCCCAGCGAGGTGGTTCACAAGTGTGTGGTCACAGTGCTCTGTGTCTTGCGGT 4500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4441 CGGGACTGCCCAGCGAGGTGGTTCACAAGTGTGTGGTCACAGTGCTCTGTGTCTTGCGGT 4500

Qy   4501 GAAGGATACCACAGTCGGCAGGTGACGTGCAAGCGGACAAAAGCCAATGGAACTGTGCAG 4560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4501 GAAGGATACCACAGTCGGCAGGTGACGTGCAAGCGGACAAAAGCCAATGGAACTGTGCAG 4560

Qy   4561 GTGGTGTCTCCAAGAGCATGTGCCCCTAAAGACCGGCCTCTGGGAAGAAAACCATGTTTT 4620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4561 GTGGTGTCTCCAAGAGCATGTGCCCCTAAAGACCGGCCTCTGGGAAGAAAACCATGTTTT 4620

Qy   4621 GGTCATCCATCTCTTCACTGGCAACCAGGCAACCGGTGTCCTGGACGTTGCATGGGCCGT 4680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4621 GGTCATCCATCTCTTCACTGGCAACCAGGCAACCGGTGTCCTGGACGTTGCATGGGCCGT 4680

Qy   4681 GCTGTGAGGATGCAGCACCGTCACACAGCTTGTCAACACAACACCTCTGACTCCAACTGT 4740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4681 GCTGTGAGGATGCAGCACCGTCACACAGCTTGTCAACACAACACCTCTGACTCCAACTGT 4740

Qy   4741 GATGACAGAAAGAGACCCACCTTAAGAAGGAACTGCACATCAGGGGCCTGTGATGTGTGT 4800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4741 GATGACAGAAAGAGACCCACCTTAAGAAGGAACTGCACATCAGGGGCCTGTGATGTGTGT 4800

Qy   4801 TGGCACACAGGCCCTTGGAAGCCCTGTACAGCAGCCTGTGGCAGGGGTTTCCAGTCTCGG 4860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4801 TGGCACACAGGCCCTTGGAAGCCCTGTACAGCAGCCTGTGGCAGGGGTTTCCAGTCTCGG 4860

Qy   4861 AAAGTCGACTGTATCCACACAAGGAGTTGCAAACCTGTGGCCAAGAGACACTGTGTACAG 4920
          ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db   4861 AAAGTCGACTGTATCCACACAACCACTTGCAAACCTGTGGCCAAGAGACACTGTGTACAG 4920

Qy   4921 AAAAAGAAACCAATTTCCTGGCGGCACTGTCTTGGGCCCTCCTGTGATAGAGACTGCACA 4980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4921 AAAAAGAAACCAATTTCCTGGCGGCACTGTCTTGGGCCCTCCTGTGATAGAGACTGCACA 4980

Qy   4981 GACACAACTCACTACTGTATGTTTGTAAAACATCTTAATTTGTGTTCTCTAGACCGCTAC 5040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4981 GACACAACTCACTACTGTATGTTTGTAAAACATCTTAATTTGTGTTCTCTAGACCGCTAC 5040

Qy   5041 AAACAAAGGTGCTGCCAGTCATGTCAAGAGGGATAA 5076
          ||||||||||||||||||||||||||||||||||||
Db   5041 AAACAAAGGTGCTGCCAGTCATGTCAAGAGGGATAA 5076
```

```
RESULT 2
AAD17757
ID   AAD17757 standard; DNA; 7260 BP.
XX
AC   AAD17757;
XX
DT   10-DEC-2001  (first entry)
XX
DE   Human novel KIAA1233 like protein, NOV 2a encoding DNA.
XX
KW   Human; NOV-X protein; KIAA1233-like protein; STE20-like protein; tumour;
KW   trypsin inhibitor-like protein; gene therapy; haematopoietic; illness;
KW   immunological disorder; neurodegenerative disorder; Alzheimer's disease;
KW   Parkinson's disease, immunomodulatory, pharmacogenomic, haemostatic;
KW   human immunodeficiency virus, HIV, fertility disorder, neuroprotective;
KW   cytostatic; nootropic; anti infertility; cancer; chromosome 15;
KW   NOV-2a protein; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   5'UTR           1..135
FT                   /*tag=  a
FT   CDS             136..5211
FT                   /*tag=  b
FT                   /product= "Human novel KIAA1233-like protein, NOV-2a"
FT   sig_peptide     136..213
FT                   /*tag=  c
FT   mat_peptide     214..5208
FT                   /*tag=  d
FT                   /product= "Human mature novel KIAA1233 like protein,
FT                   NOV-2a"
FT   3'UTR           5212..7260
FT                   /*tag=  e
XX
PN   WO200162928-A2.
XX
PD   30-AUG-2001.
XX
PF   26-FEB-2001; 2001WO-US06151.
XX
PR   25-FEB-2000; 2000US-0184951.
PR   28-FEB-2000; 2000US-0185548.
PR   01-MAR-2000; 2000US-0185967.
PR   18-APR-2000; 2000US-0197723.
PR   27-APR-2000; 2000US-0199957.
PR   23-FEB-2001; 2001US-0789390.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Vernet CAM,  Fernandes E,  Shimkets RA,  Macdougall J,  Spaderna SK;
XX
DR   WPI; 2001-582051/65.
DR   P-PSDB; AAE10609.
XX
PT   New isolated KIAA1233 like, STE20 like, or trypsin inhibitor like
PT   polypeptide for diagnosing and treating pathological disorders, such as
PT   Parkinson's disease and for use in pharmacogenomics -
XX
PS   Claim 9; Page 20-22; 189pp; English.
XX
CC   The invention relates to novel human polypeptides referred as NOV-X
CC   and their corresponding nucleic acid sequences. NOV-X collectively
CC   include NOV-1, NOV-2a and NOV 2b which are novel KIAA1233 like
CC   polypeptides, NOV 3a, NOV 3b, NOV 3c and NOV 3d which are novel
CC   STE20 like polypeptides and NOV 4a, NOV 4b, NOV 4c, NOV 4d and NOV-4e
CC   which are novel trypsin inhibitor-like polypeptides. NOV X is used to
CC   identify a potential therapeutic agent that can modulate its activity
CC   and can be used for treating a pathology related to aberrant expression
CC   or aberrant physiological interactions of NOV-X. NOV-X or its DNA is
CC   used to determine the presence or predisposition to a disease associated
CC   with altered levels of NOV X. NOV X, its DNA and its antibody are used
CC   to treat or prevent a pathology associated with NOV X. The pathological
CC   states that can be treated or prevented are haematopoietic, cancer,
CC   immunological, tumour, neurodegenerative (e.g. Alzheimer's and
CC   Parkinson's disease), human immunodeficiency virus (HIV) illness and
CC   fertility disorders. NOV X and its DNA are used in pharmacogenomics for
CC   predictive medicine. NOV X DNA is used in gene therapy. The present
CC   sequence is a DNA encoding human novel KIAA1233 like protein, NOV-2a.
CC   NOV 2a gene is located at chromosome 15.
XX
SQ   Sequence 7260 BP; 2022 A; 1593 C; 1811 G; 1834 T; 0 other;

  Query Match             99.9%;  Score 5069.6;  DB 22;  Length 7260;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 5072;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 ATGGCTTCCTGGACGAGCCATGGTGGGTGCTGATAGGGATGGTCTTCATGCACTCTCCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        136 ATGGCTTCCTGGACGAGCCATGGTGGGTGCTGATAGGGATGGTCTTCATGCACTCTCCC 195

Qy         61 CTCCCGCAGACCACAGCTGAGAAATCTCCTGGAGCCTATTTCCTTCCCGAGTTTGCACTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        196 CTCCCGCAGACCACAGCTGAGAAATCTCCTGGAGCCTATTTCCTTCCCGAGTTTGCACTT 255

Qy        121 TCTCCTCAGGGAAGTTTTCTGGAAGACACAACAGGGGAGCAGTTCCTCACTTATCGCTAT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        256 TCTCCTCAGGGAAGTTTTCTGGAAGACACAACAGGGGAGCAGTTCCTCACTTATCGCTAT 315

Qy        181 GATGACCAGACCTCAAGAAACACTCGTTCAGATGAAGACAAAGATGGCAACTGGGATGCT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        316 GATGACCAGACCTCAAGAAACACTCGTTCAGATGAAGACAAAGATGGCAACTGGGATGCT 375

Qy        241 TGGGGCGACTGGAGTGACTGCTCCCGGACCTGTGGGGGAGGAGCATCATATTCTCTGCGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        376 TGGGGCGACTGGAGTGACTGCTCCCGGACCTGTGGGGGAGGAGCATCATATTCTCTGCGG 435

Qy        301 AGATGTTTGACTGGAAGGAATTGTGAAGGGCAGAACATTCGGTACAAGACATGCAGCAAT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        436 AGATGTTTGACTGGAAGGAATTGTGAAGGGCAGAACATTCGGTACAAGACATGCAGCAAT 495

Qy        361 CATGACTGCCCTCCAGATGGAGAAGATTTCAGAGCCCAGCAGTGCTCAGCCTACAATGAT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        496 CATGACTGCCCTCCAGATGGAGAAGATTTCAGAGCCCAGCAGTGCTCAGCCTACAATGAT 555

Qy        421 GTCCAGTATCAGGGGCATTACTATGAATGGCTTCCACGATATAATGATCCTGCTGCCCCC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        556 GTCCAGTATCAGGGGCATTACTATGAATGGCTTCCACGATATAATGATCCTGCTGCCCCC 615

Qy        481 TGTGCACTCAAGTGTCATGCACAAGGACAAAACTTGGTGGTGGAGCTGGCACCTAAGGTA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        616 TGTGCACTCAAGTGTCATGCACAAGGACAAAACTTGGTGGTGGAGCTGGCACCTAAGGTA 675

Qy        541 CTGGATGGAACTCGTTGCAACACGGACTCCTTGGACATGTGTATCAGTGGCATCTGTCAG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        676 CTGGATGGAACTCGTTGCAACACGGACTCCTTGGACATGTGTATCAGTGGCATCTGTCAG 735

Qy        601 GCAGTGGGCTGCGATCGGCAACTGGGAAGCAATGCCAAGGAGGACAACTGTGGAGTCTGT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        736 GCAGTGGGCTGCGATCGGCAACTGGGAAGCAATGCCAAGGAGGACAACTGTGGAGTCTGT 795

Qy        661 GCCGGCGATGGCTCCACCTGCAGGCTTGTACGGGGACAATCAAAGTCACACGTTTCTCCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        796 GCCGGCGATGGCTCCACCTGCAGGCTTGTACGGGGACAATCAAAGTCACACGTTTCTCCT 855

Qy        721 GAAAAAAGAGAAGAAAATGTAATTGCTGTTCCTTCGGGAAGTCGAAGTGTGAGAATTACA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        856 GAAAAAAGAGAAGAAAATGTAATTGCTGTTCCTTCGGGAAGTCGAAGTGTGAGAATTACA 915

Qy        781 GTGAAAGGACCTGCCCACCTCTTTATTGAATCAAAAACACTTCAAGGAAGCAAAGGAGAA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        916 GTGAAAGGACCTGCCCACCTCTTTATTGAATCAAAAACACTTCAAGGAAGCAAAGGAGAA 975
```

```
Qy        841 CACAGCTTTAACAGCCCCGGCGTCTTTGTCGTAGAAAACACAACAGTGGAATTCCAGAGG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        976 CACAGCTTTAACAGCCCCGGCGTCTTTGTCGTAGAAAACACAACAGTGGAATTCCAGAGG 1035

Qy        901 GGCTCCGAGAGGCAAACTTTTAAGATTCCAGGACCTCTGATGGCTGATTTCATCTTCAAG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1036 GGCTCCGAGAGGCAAACTTTTAAGATTCCAGGACCTCTGATGGCTGATTTCATCTTCAAG 1095

Qy        961 ACCAGGTACACTGCAGCCAAAGACAGCGTGGTTCAGTTCTTCTTTTACCAGCCCATCAGT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1096 ACCAGGTACACTGCAGCCAAAGACAGCGTGGTTCAGTTCTTCTTTTACCAGCCCATCAGT 1155

Qy       1021 CATCAGTGGAGACAAACTGACTTCTTTCCCTGCACTGTGACGTGTGGAGGAGGTTATCAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1156 CATCAGTGGAGACAAACTGACTTCTTTCCCTGCACTGTGACGTGTGGAGGAGGTTATCAG 1215

Qy       1081 CTCAATTCTGCTGAATGTGTGGATATCCGCTTGAAGAGGGTAGTTCCTGACCATTATTGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1216 CTCAATTCTGCTGAATGTGTGGATATCCGCTTGAAGAGGGTAGTTCCTGACCATTATTGT 1275

Qy       1141 CACTACTACCCTGAAAATGTAAAACCAAAACCAAAACTGAAGGAATGCAGCATGGATCCC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1276 CACTACTACCCTGAAAATGTAAAACCAAAACCAAAACTGAAGGAATGCAGCATGGATCCC 1335

Qy       1201 TGCCCATCAAGTGATGGATTTAAAGAGATAATGCCCTATGACCACTTCCAACCTCTTCCT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1336 TGCCCATCAAGTGATGGATTTAAAGAGATAATGCCCTATGACCACTTCCAACCTCTTCCT 1395

Qy       1261 CGCTGGGAACATAATCCTTGGACTGCATGTTCCGTGTCCTGTGGAGGAGGGATTCAGAGA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1396 CGCTGGGAACATAATCCTTGGACTGCATGTTCCGTGTCCTGTGGAGGAGGGATTCAGAGA 1455

Qy       1321 CGGAGCTTTGTGTGTGTAGAGGAATCCATGCATGGAGAGATATTGCAGGTGGAAGAATGG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1456 CGGAGCTTTGTGTGTGTAGAGGAATCCATGCATGGAGAGATATTGCAGGTGGAAGAATGG 1515

Qy       1381 AAGTGCATGTACGCACCCAAACCCAAGGTTATGCAAACTTGTAATCTGTTTGATTGCCCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1516 AAGTGCATGTACGCACCCAAACCCAAGGTTATGCAAACTTGTAATCTGTTTGATTGCCCC 1575

Qy       1441 AAGTGGATTGCCATGGAGTGGTCTCAGTGCACAGTGACTTGTGGCCGAGGGTTACGGTAC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1576 AAGTGGATTGCCATGGAGTGGTCTCAGTGCACAGTGACTTGTGGCCGAGGGTTACGGTAC 1635

Qy       1501 CGGGTTGTTCTGTGTATTAACCACCGCGGAGAGCATGTTGGGGGCTGCAATCCACAACTG 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1636 CGGGTTGTTCTGTGTATTAACCACCGCGGAGAGCATGTTGGGGGCTGCAATCCACAACTG 1695

Qy       1561 AAGTTACACATCAAAGAAGAATGTGTCATTCCCATCCCGTGTTATAAACCAAAAGAAAAA 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1696 AAGTTACACATCAAAGAAGAATGTGTCATTCCCATCCCGTGTTATAAACCAAAAGAAAAA 1755

Qy       1621 AGTCCAGTGGAAGCAAAATTGCCTTGGCTGAAACAAGCACAAGAACTAGAAGAGACCAGA 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1756 AGTCCAGTGGAAGCAAAATTGCCTTGGCTGAAACAAGCACAAGAACTAGAAGAGACCAGA 1815

Qy       1681 ATAGCAACAGAAGAACCAACGTTCATTCCAGAACCCTGGTCAGCTTGCAGTACTACGTGT 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1816 ATAGCAACAGAAGAACCAACGTTCATTCCAGAACCCTGGTCAGCTTGCAGTACTACGTGT 1875

Qy       1741 GGGCCAGGTGTGCAGGTCCGCGAGGTGAAGTGCCGTGTGCTCCTCACATTCACGCAGACT 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1876 GGGCCAGGTGTGCAGGTCCGCGAGGTGAAGTGCCGTGTGCTCCTCACATTCACGCAGACT 1935

Qy       1801 GAGACTGAGCTGCCCGAGGAAGAGTGTGAAGGCCCCAAGCTGCCCACCGAACGGCCCTGC 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1936 GAGACTGAGCTGCCCGAGGAAGAGTGTGAAGGCCCCAAGCTGCCCACCGAACGGCCCTGC 1995

Qy       1861 CTCCTGGAAGCATGTGATGAGAGCCCGGCCTCCCGAGAGCTAGACATCCCTCTCCCTGAG 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1996 CTCCTGGAAGCATGTGATGAGAGCCCGGCCTCCCGAGAGCTAGACATCCCTCTCCCTGAG 2055

Qy       1921 GACAGTGAGACGACTTACGACTGGGAGTACGCTGGGTTCACCCCTTGCACAGCAACATGT 1980
```

```
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2056 GACAGTGAGACGACTTACGACTGGGAGTACGCTGGGTTCACCCCTTGCACAGCAACATGC 2115

Qy     1981 TTGGGAGGCCATCAAGAAGCCATAGCAGTGTGCTTACATATCCAGACCCAGCAGACAGTC 2040
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2116 GTGGGAGGCCATCAAGAAGCCATAGCAGTGTGCTTACATATCCAGACCCAGCAGACAGTC 2175

Qy     2041 AATGACAGCTTGTGTGATATGGTCCACCGTCCTCCAGCCATGAGCCAGGCCTGTAACACA 2100
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2176 AATGACAGCTTGTGTGATATGGTCCACCGTCCTCCAGCCATGAGCCAGGCCTGTAACACA 2235

Qy     2101 GAGCCCTGTCCCCCCAGGTGGCATGTGGGCTCTTGGGGGCCCTGCTCAGCTACCTGTGGA 2160
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2236 GAGCCCTGTCCCCCCAGGTGGCATGTGGGCTCTTGGGGGCCCTGCTCAGCTACCTGTGGA 2295

Qy     2161 GTTGGAATTCAGACCCGAGATGTGTACTGCCTGCACCTAGGGGAGACCCCTGCCCCTCCT 2220
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2296 GTTGGAATTCAGACCCGAGATGTGTACTGCCTGCACCTAGGGGAGACCCCTGCCCCTCCT 2355

Qy     2221 GAGGAGTGCCGAGATGAAAAGCCCCATGCTTTACAAGCATGCAATCAGTTTGACTGCCCT 2280
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2356 GAGGAGTGCCGAGATGAAAAGCCCCATGCTTTACAAGCATGCAATCAGTTTGACTGCCCT 2415

Qy     2281 CCTGGCTGGCACATTGAAGAATGGCAGCAGTGTTCCAGGACTTGTGGCGGGGGAACTCAG 2340
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2416 CCTGGCTGGCACATTGAAGAATGGCAGCAGTGTTCCAGGACTTGTGGCGGGGGAACTCAG 2475

Qy     2341 AACAGAAGAGTCACCTGTCGGCAGCTGCTAACGGATGGCAGCTTTTTGAATCTCTCAGAT 2400
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2476 AACAGAAGAGTCACCTGTCGGCAGCTGCTAACGGATGGCAGCTTTTTGAATCTCTCAGAT 2535

Qy     2401 GAATTGTGCCAAGGACCCAAGGCATCGTCTCACAAGTCCTGTGCCAGGACAGACTGTCCT 2460
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2536 GAATTGTGCCAAGGACCCAAGGCATCGTCTCACAAGTCCTGTGCCAGGACAGACTGTCCT 2595

Qy     2461 CCACATTTAGCTGTGGGAGACTGGTCGAAGTGTTCTGTCAGTTGTGGTGTTGGAATCCAG 2520
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2596 CCACATTTAGCTGTGGGAGACTGGTCGAAGTGTTCTGTCAGTTGTGGTGTTGGAATCCAG 2655

Qy     2521 AGAAGAAAGCAGGTGTGTCAAAGGCTGGCAGCCAAAGGTCGGCGCATCCCCCTCAGTGAG 2580
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2656 AGAAGAAAGCAGGTGTGTCAAAGGCTGGCAGCCAAAGGTCGGCGCATCCCCCTCAGTGAG 2715

Qy     2581 ATGATGTGCAGGGATCTACCAGGGTTCCCTCTTGTAAGATCTTGCCAGATGCCTGAGTGC 2640
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2716 ATGATGTGCAGGGATCTACCAGGGTTCCCTCTTGTAAGATCTTGCCAGATGCCTGAGTGC 2775

Qy     2641 AGTAAAATCAAATCAGAGATGAAGACAAAACTTGGTGAGCAGGGTCCGCAGATCCTCAGT 2700
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2776 AGTAAAATCAAATCAGAGATGAAGACAAAACTTGGTGAGCAGGGTCCGCAGATCCTCAGT 2835

Qy     2701 GTCCAGAGAGTCTACATTCAGACAAGGGAAGAGAAGCGTATTAACCTGACCATTGGTAGC 2760
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2836 GTCCAGAGAGTCTACATTCAGACAAGGGAAGAGAAGCGTATTAACCTGACCATTGGTAGC 2895

Qy     2761 AGAGCCTATTTGCTGCCCAACACATCCGTGATTATTAAGTGCCCCGTGCGACGATTCCAG 2820
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2896 AGAGCCTATTTGCTGCCCAACACATCCGTGATTATTAAGTGCCCCGTGCGACGATTCCAG 2955

Qy     2821 AAATCTCTGATCCAGTGGGAGAAGGATGGCCGTTGCCTGCAGAACTCCAAACGGCTTGGC 2880
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2956 AAATCTCTGATCCAGTGGGAGAAGGATGGCCGTTGCCTGCAGAACTCCAAACGGCTTGGC 3015

Qy     2881 ATCACCAAGTCAGGCTCACTAAAAATCCACGGTCTTGCTGCCCCCGACATCGGCGTGTAC 2940
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3016 ATCACCAAGTCAGGCTCACTAAAAATCCACGGTCTTGCTGCCCCCGACATCGGCGTGTAC 3075

Qy     2941 CGGTGCATTGCAGGCTCTGCACAGGAAACAGTTGTGCTCAAGCTCATTGGTACTGACAAC 3000
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3076 CGGTGCATTGCAGGCTCTGCACAGGAAACAGTTGTGCTCAAGCTCATTGGTACTGACAAC 3135

Qy     3001 CGGCTCATCGCACGCCCAGCCCTCAGGGAGCCTATGAGGGAATATCCTGGGATGGACCAC 3060
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3136 CGGCTCATCGCACGCCCAGCCCTCAGGGAGCCTATGAGGGAATATCCTGGGATGGACCAC 3195

Qy     3061 AGCGAAGCCAATAGTTTGGGAGTCACATGGCACAAAATGAGGCAAATGTGGAATAACAAA 3120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3196 AGCGAAGCCAATAGTTTGGGAGTCACATGGCACAAAATGAGGCAAATGTGGAATAACAAA 3255

Qy     3121 AATGACCTTTATCTGGATGATGACCACATTAGTAACCAGCCTTTCTTGAGAGCTCTGTTA 3180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3256 AATGACCTTTATCTGGATGATGACCACATTAGTAACCAGCCTTTCTTGAGAGCTCTGTTA 3315

Qy     3181 GGCCACTGCAGCAATTCTGCAGGAAGCACCAACTCCTGGGAGTTGAAGAATAAGCAGTTT 3240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3316 GGCCACTGCAGCAATTCTGCAGGAAGCACCAACTCCTGGGAGTTGAAGAATAAGCAGTTT 3375

Qy     3241 GAAGCAGCAGTTAAACAAGGAGCATATAGCATGGATACAGCCCAGTTTGATGAGCTGATA 3300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3376 GAAGCAGCAGTTAAACAAGGAGCATATAGCATGGATACAGCCCAGTTTGATGAGCTGATA 3435

Qy     3301 AGAAACATGAGTCAGCTCATGGAAACCGGAGAGGTCAGCGATGATCTTGCGTCCCAGCTG 3360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3436 AGAAACATGAGTCAGCTCATGGAAACCGGAGAGGTCAGCGATGATCTTGCGTCCCAGCTG 3495

Qy     3361 ATATATCAGCTGGTGGCCGAATTAGCCAAGGCACAGCCAACACACATGCAGTGGCGGGGC 3420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3496 ATATATCAGCTGGTGGCCGAATTAGCCAAGGCACAGCCAACACACATGCAGTGGCGGGGC 3555

Qy     3421 ATCCAGGAAGAGACACCTCCTGCTGCTCAGCTCAGAGGGGAAACAGGGAGTGTGTCCCAA 3480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3556 ATCCAGGAAGAGACACCTCCTGCTGCTCAGCTCAGAGGGGAAACAGGGAGTGTGTCCCAA 3615

Qy     3481 AGCTCGCATGCAAAAAACTCAGGCAAGCTGACATTCAAGCCGAAAGGACCTGTTCTCATG 3540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3616 AGCTCGCATGCAAAAAACTCAGGCAAGCTGACATTCAAGCCGAAAGGACCTGTTCTCATG 3675

Qy     3541 AGGCAAAGCCAACCTCCCTCAATTTCATTTAATAAAACAATAAATTCCAGGATTGGAAAT 3600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3676 AGGCAAAGCCAACCTCCCTCAATTTCATTTAATAAAACAATAAATTCCAGGATTGGAAAT 3735

Qy     3601 ACAGTATACATTACAAAAAGGACAGAGGTCATCAATATACTGTGTGACCTTATTACCCCC 3660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3736 ACAGTATACATTACAAAAAGGACAGAGGTCATCAATATACTGTGTGACCTTATTACCCCC 3795

Qy     3661 AGTGAGGCCACATATACATGGACCAAGGATGGAACCTTGTTACAGCCCTCAGTAAAAATA 3720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3796 AGTGAGGCCACATATACATGGACCAAGGATGGAACCTTGTTACAGCCCTCAGTAAAAATA 3855

Qy     3721 ATTTTGGATGGAACTGGGAAGATACAGATACAGAATCCTACAAGGAAAGAACAAGGCATA 3780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3856 ATTTTGGATGGAACTGGGAAGATACAGATACAGAATCCTACAAGGAAAGAACAAGGCATA 3915

Qy     3781 TATGAATGTTCTGTAGCTAATCATCTTGGTTCAGATGTGGAAAGTTCTTCTGTGCTGTAT 3840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3916 TATGAATGTTCTGTAGCTAATCATCTTGGTTCAGATGTGGAAAGTTCTTCTGTGCTGTAT 3975

Qy     3841 GCAGAGGCACCTGTCATCTTGTCTGTTGAAAGAAATATCACCAAACCAGAGCACAACCAT 3900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3976 GCAGAGGCACCTGTCATCTTGTCTGTTGAAAGAAATATCACCAAACCAGAGCACAACCAT 4035

Qy     3901 CTGTCTGTTGTGGTTGGAGGCATCGTGGAGGCAGCCCTTGGAGCAAACGTGACAATCCGA 3960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4036 CTGTCTGTTGTGGTTGGAGGCATCGTGGAGGCAGCCCTTGGAGCAAACGTGACAATCCGA 4095

Qy     3961 TGTCCTGTAAAAGGTGTCCCTCAGCCTAATATAACTTGGTTGAAGAGAGGAGGATCTCTG 4020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4096 TGTCCTGTAAAAGGTGTCCCTCAGCCTAATATAACTTGGTTGAAGAGAGGAGGATCTCTG 4155

Qy     4021 AGTGGCAATGTTTCCTTGCTTTTCAATGGATCCCTGTTGTTGCAGAATGTTTCCCTTGAA 4080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4156 AGTGGCAATGTTTCCTTGCTTTTCAATGGATCCCTGTTGTTGCAGAATGTTTCCCTTGAA 4215

Qy     4081 AATGAAGGAACCTACGTCTGCATAGCCACCAATGCTCTTGGAAAGGCAGTGGCAACATCT 4140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4216 AATGAAGGAACCTACGTCTGCATAGCCACCAATGCTCTTGGAAAGGCAGTGGCAACATCT 4275
```

```
Qy       4141 GTACTCCACTTGCTGGAACGAAGATGGCCAGAGAGTAGAATCGTATTTCTGCAAGGACAT 4200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4276 GTACTCCACTTGCTGGAACGAAGATGGCCAGAGAGTAGAATCGTATTTCTGCAAGGACAT 4335

Qy       4201 AAAAAGTACATTCTTCTAGCTAACTAACACTAGAACTAACAGCAATGACCTAACAGGAGAA 4260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4336 AAAAAGTACATTCTTCTAGCTAACTAACACTAGAACTAACAGCAATGACCTAACAGGAGAA 4395

Qy       4261 CCCCTGCCTCAAGACCCTTTTTCGGAGCCTGGTAACTGGTCACATTGTTCTGCCACCTGT 4320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4396 CCCCTGCCTCAAGACCCTTTTTCGGAGCCTGGTAACTGGTCACATTGTTCTGCCACCTGT 4455

Qy       4321 GGTCATTTGGGAGCCCGCATTCAGAGACCCCAGTGTGTGATGGCCAATGGGCAGGAAGTG 4380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4456 GGTCATTTGGGAGCCCGCATTCAGAGACCCCAGTGTGTGATGGCCAATGGGCAGGAAGTG 4515

Qy       4381 AGTGACCCCTGTGTGATCACTCCCAGAAGCCACTGGCTGGGTTTGAGCCCTGTAACATC 4440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4516 AGTGACCCCTGTGTGATCACTCCCAGAAGCCACTGGCTGGGTTTGAGCCCTGTAACATC 4575

Qy       4441 CGGGACTGCCCAGGCAGGTGGTTCACAAGTGTGTGGTCACAGTGCTCTGTGCTTGCCGGT 4500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4576 CGGGACTGCCCAGGCAGGTGGTTCACAAGTGTGTGGTCACAGTGCTCTGTGCTTGCCGGT 4635

Qy       4501 GAATGATACCACAGTGCCTACTTGACTGTCAAGCGACAAAAGCCAATCGAACCTGTCATC 4560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4636 GAATGATACCACAGTGCCTACTTGACTGTCAAGCGACAAAAGCCAATCGAACCTGTCATC 4695

Qy       4561 GTGGTGTCTCCAAGAGCATGTGCCCTAAAGACCGGCCTCTGGGAAGAAAACCATGCTTT 4620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4696 GTGGTGTCTCCAAGAGCATGTGCCCTAAAGACCGGCCTCTGGGAAGAAAACCATGCTTT 4755

Qy       4621 GGTCATCCATGTGTTCAGTGGGAACCAGGGAACCGGTGTCCTGGACGTTGCATGCGCCGT 4680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4756 GGTCATCCATGTGTTCAGTGGGAACCAGGGAACCGGTGTCCTGGACGTTGCATGCGCCGT 4815

Qy       4681 GCTGTGAGGATGCAGCAGCGTCACACAGCTTGTCAACACAACAGCTCTGACTCCAACTGT 4740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4816 GCTGTGAGGATGCAGCAGCGTCACACAGCTTGTCAACACAACAGCTCTGACTCCAACTGT 4875

Qy       4741 GATGACAGAAAGAGACCCACCTTAAGAAGGAACTGCACATCAGGGGCCTGTGATGTGTGT 4800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4876 GATGACAGAAAGAGACCCACCTTAAGAAGGAACTGCACATCAGGGGCCTGTGATGTGTGT 4935

Qy       4801 TGGCACACAGGCCCTTGGAAGCCCTGTACAGCAGCCTGTGGGAGGGGTTCCAGTCTGCC 4860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4936 TGGCACACAGGCCCTTGGAAGCCCTGTACAGCAGCCTGTGGGAGGGGTTCCAGTCTGCC 4995

Qy       4861 AAAGTCGACTGTATCCACACAAGGAGTTGCAAACCTGTGGCCAAGAGACACTGTGTACAG 4920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4996 AAAGTCGACTGTATCCACACAAGGAGTTGCAAACCTGTGGCCAAGAGACACTGTGTACAG 5055

Qy       4921 AAAAAGAAACCAATTTCCTGGCGGCACTGTCTTGGGCCTCCTGTGATAGAGACTGCACA 4980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5056 AAAAAGAAACCAATTTCCTGGCGGCACTGTCTTGGGCCTCCTGTGATAGAGACTGCACA 5115

Qy       4981 GACACAACTCACTACTGTATGTTTGTAAAACATCTTAATTTGTGTTCTCTAGACCGCTAC 5040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5116 GACACAACTCACTACTGTATGTTTGTAAAACATCTTAATTTGTGTTCTCTAGACCGCTAC 5175

Qy       5041 AAACAAAGGTGCTGCCAGTCATGTCAAGAGGGATAA 5076
              ||||||||||||||||||||||||||||||||||||
Db       5176 AAACAAAGGTGCTGCCAGTCATGTCAAGAGGGATAA 5211


RESULT 3
AAD14371
ID   AAD14371 standard; cDNA; 4854 BP.
XX
AC   AAD14371;
XX
DT   01-NOV-2001  (first entry)
XX
DE   Novel human protein (NHP) cDNA #8.
XX
KW   Novel human protein; NHP; thrombospondin; gene therapy; cancer;
KW   cytostatic; antisense therapy; angiogenesis; biological disorder; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..4854
FT                   /*tag=  a
FT                   /product= "Novel human protein"
XX
PN   WO200161011-A2.
XX
PD   23-AUG-2001.
XX
PF   15-FEB-2001; 2001WO-US05290.
XX
PR   17-FEB-2000; 2000US-0183282.
XX
PA   (LEXI-) LEXICON GENETICS INC.
XX
PI   Donoho G,  Scoville J,  Turner CA,  Friedrich G,  Zambrowicz B;
PI   Sands AT;
XX
DR   WPI; 2001-514776/56.
DR   P-PSDB; AAE07870.
XX
PT   An isolated nucleic acid encoding a thrombospondin, useful as a
PT   hybridization probe and gene therapy treatments of cancer
XX
PS   Claim 4; Page 48-49; 56pp; English.
XX
CC   The invention relates to novel human polynucleotides encoding proteins
CC   that share sequence similarity with animal proteins having thrombospondin
CC   repeats. NHPs are expressed in human cell lines, pituitary, lymph nodes,
CC   prostrate, testis, adrenal gland, uterus, foetal kidney, foetal lung and
CC   gene trapped human cells. NHPs are thrombospondins useful for treating
CC   biological disorders involving angiogenesis, cancer and development and
CC   also in pharmacogenomic applications. NHPs are useful as a hybridisation
CC   probe for screening libraries, assessing gene expression patterns and
CC   also in gene therapy. Proteins having thrombospondin repeats act as
CC   receptors, secreted extracellular matrix proteins and proteases. The
CC   present sequence is a cDNA encoding novel human protein (NHP).
XX
SQ   Sequence 4854 BP; 1331 A; 1175 C; 1276 G; 1072 T; 0 other;

  Query Match             93.9%;  Score 4766;  DB 22;  Length 4854;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 4766;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        311 CTGGAAGGAATTGTGAAGGGAGAACATTCGGTACAAGACATGCAGCAATCATGACTGCC 370
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         89 CTGGAAGGAATTGTGAAGGGAGAACATTCGGTACAAGACATGCAGCAATCATGACTGCC 148

Qy        371 CTCCAGATGCAGAAGATTTCAGAGCCAGCAGGTCTAGCCACAATGATGCCAGTATC 430
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        149 CTCCAGATGCAGAAGATTTCAGAGCCAGCAGGTCTAGCCACAATGATGCCAGTATC 208

Qy        431 AGGGGCATTACTATGAATGGCTTCCACGATATAATGATCCTGCTGCCCCGTGTGCACTCA 490
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        209 AGGGGCATTACTATGAATGGCTTCCACGATATAATGATCCTGCTGCCCCGTGTGCACTCA 268

Qy        491 AGTGTCATGCACAAGGACAAAACTTGGTGGTGGAGCTGGCACCTAAGGTACTGGATGGAA 550
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        269 AGTGTCATGCACAAGGACAAAACTTGGTGGTGGAGCTGGCACCTAAGGTACTGGATGGAA 328

Qy        551 CTCGTTGCAACACGGACTTCCTGGACATGTGTATCAGTGGCATCGTCAGCAGTGGGCT 610
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        329 CTCGTTGCAACACGGACTTCCTGGACATGTGTATCAGTGGCATCGTCAGCAGTGGGCT 388

Qy        611 GCGATGGGCAACTGGGAAGAAGGCAAGGAGGACAACTGCTGAGTCTGTGCCGCATG 670
```

```
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        389 GCGATCGGCAACTGGGAAGCAATGCCAAGGAGGACAACTGTGGAGTCTGTGCCGGCGATG 448

Qy        671 GCTCCACCTGCAGGCTTGTACGGGGACAATCAAAGTCACACGTTTCTCCTGAAAAAAGAG 730
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        449 GCTCCACCTGCAGGCTTGTACGGGGACAATCAAAGTCACACGTTTCTCCTGAAAAAAGAG 508

Qy        731 AAGAAAATGTAATTGCTGTTCCTTTGGGAAGTCGAAGTGTGAGAATTACAGTGAAAGGAC 790
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        509 AAGAAAATGTAATTGCTGTTCCTTTGGGAAGTCGAAGTGTGAGAATTACAGTGAAAGGAC 568

Qy        791 CTGCCCACCTCTTTATTGAATCAAAAACACTTCAAGGAAGCAAAGGAGAACACAGCTTTA 850
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        569 CTGCCCACCTCTTTATTGAATCAAAAACACTTCAAGGAAGCAAAGGAGAACACAGCTTTA 628

Qy        851 ACAGCCCCGGCGTCTTTGTCGTAGAAAACACAACAGTGGAATTTCAGAGGGGCTCCGAGA 910
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        629 ACAGCCCCGGCGTCTTTGTCGTAGAAAACACAACAGTGGAATTTCAGAGGGGCTCCGAGA 688

Qy        911 GGCAAACTTTTAAGATTCCAGGACCTCTGATGGCTGATTTCATCTTCAAGACCAGGTACA 970
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        689 GGCAAACTTTTAAGATTCCAGGACCTCTGATGGCTGATTTCATCTTCAAGACCAGGTACA 748

Qy        971 CTGCAGCCAAAGACAGCGTGGTTCAGTTCTTCTTTTACCAGCCCATCAGTCATCAGTGGA 1030
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        749 CTGCAGCCAAAGACAGCGTGGTTCAGTTCTTCTTTTACCAGCCCATCAGTCATCAGTGGA 808

Qy       1031 GACAAACTGACTTCTTTCCCTGCACTGTGACGTGTGGAGGAGGTTATCAGCTCAATTCTG 1090
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        809 GACAAACTGACTTCTTTCCCTGCACTGTGACGTGTGGAGGAGGTTATCAGCTCAATTCTG 868

Qy       1091 CTGAATGTGTGGATATCCGCTTGAAGAGGGTAGTTCCTGACCATTATTGTCACTACTACC 1150
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        869 CTGAATGTGTGGATATCCGCTTGAAGAGGGTAGTTCCTGACCATTATTGTCACTACTACC 928

Qy       1151 CTGAAAATGTAAAACCAAAACCAAAACTGAAGGAATGCAGCATGGATCCCTGCCCATCAA 1210
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        929 CTGAAAATGTAAAACCAAAACCAAAACTGAAGGAATGCAGCATGGATCCCTGCCCATCAA 988

Qy       1211 GTGATGGATTTAAAGAGATAATGCCCTATGACCACTTCCAACCTCTTCCTCGCTGGGAAC 1270
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        989 GTGATGGATTTAAAGAGATAATGCCCTATGACCACTTCCAACCTCTTCCTCGCTGGGAAC 1048

Qy       1271 ATAATCCTTGGACTGCATGTTCCGTGTCCTGTGGAGGAGGGATTCAGAGACGGAGCTTTG 1330
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1049 ATAATCCTTGGACTGCATGTTCCGTGTCCTGTGGAGGAGGGATTCAGAGACGGAGCTTTG 1108

Qy       1331 TGTGTGTAGAGGAATCCATGCATGGAGAGATATTGCAGGTGGAAGAATGGAAGTGCATGT 1390
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1109 TGTGTGTAGAGGAATCCATGCATGGAGAGATATTGCAGGTGGAAGAATGGAAGTGCATGT 1168

Qy       1391 ACGCACCCAAACCCAAGGTTATGCAAACTTGTAATCTGTTTGATTGCCCCAAGTGGATTG 1450
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1169 ACGCACCCAAACCCAAGGTTATGCAAACTTGTAATCTGTTTGATTGCCCCAAGTGGATTG 1228

Qy       1451 CCATGGAGTGGTCTCAGTGCACAGTGACTTGTGGCCCAGGGTTACGGTACCGGGTTGTTC 1510
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1229 CCATGGAGTGGTCTCAGTGCACAGTGACTTGTGGCCCAGGGTTACGGTACCGGGTTGTTC 1288

Qy       1511 TGTGTATTAACCACCGCGGAGAGCATGTTGGGGGCTGCAATCCACAACTGAAGTTACACA 1570
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1289 TGTGTATTAACCACCGCGGAGAGCATGTTGGGGGCTGCAATCCACAACTGAAGTTACACA 1348

Qy       1571 TCAAAGAAGAATGTGTCATTCCCATCCCGTGTTATAAACCAAAAGAAAAAAGTCCAGTGG 1630
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1349 TCAAAGAAGAATGTGTCATTCCCATCCCGTGTTATAAACCAAAAGAAAAAAGTCCAGTGG 1408

Qy       1631 AAGCAAAATTGCCTTGGCTGAAACAAGCACAAGAACTAGAAGAGACCAGAATAGCAACAG 1690
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1409 AAGCAAAATTGCCTTGGCTGAAACAAGCACAAGAACTAGAAGAGACCAGAATAGCAACAG 1468

Qy       1691 AAGAACCAACGTTCATTCCAGAACCCTGGTCAGCCTGCAGTACCACGTGTGGGCCAGGTG 1750
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1469 AAGAACCAACGTTCATTCCAGAACCCTGGTCAGCCTGCAGTACCACGTGTGGGCCAGGTG 1528

Qy       1751 TGCAGGTCCGCGAGGTGAAGTGCCGTGTGCTCCTCACATTCACGCAGACTGAGACTGAGC 1810
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1529 TGCAGGTCCGCGAGGTGAAGTGCCGTGTGCTCCTCACATTCACGCAGACTGAGACTGAGC 1588

Qy       1811 TGCCCGAGGAAGAGTGTGAAGGCCCCAAGCTGCCCACCGAACGGCCCTGCCTCCTGGAAG 1870
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1589 TGCCCGAGGAAGAGTGTGAAGGCCCCAAGCTGCCCACCGAACGGCCCTGCCTCCTGGAAG 1648

Qy       1871 CATGTGATGAGAGCCCGGCCTCCCGAGAGCTAGACATCCCTCTCCCTGAGGACAGTGAGA 1930
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1649 CATGTGATGAGAGCCCGGCCTCCCGAGAGCTAGACATCCCTCTCCCTGAGGACAGTGAGA 1708

Qy       1931 CGACTTACGACTGGGAGTACGCTGGGTTCACCCCTTGCACAGCAACATGCTTGGGAGGCC 1990
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1709 CGACTTACGACTGGGAGTACGCTGGGTTCACCCCTTGCACAGCAACATGCTTGGGAGGCC 1768

Qy       1991 ATCAAGAAGCCATAGCAGTGTGCTTACATATCCAGACCCAGCAGACACTCAATGACAGCT 2050
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1769 ATCAAGAAGCCATAGCAGTGTGCTTACATATCCAGACCCAGCAGACACTCAATGACAGCT 1828

Qy       2051 TGTGTGATATGGTCCACCGTCCTCCAGCCATGAGCCAGGCCTGTAACACAGAGCCCTGTC 2110
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1829 TGTGTGATATGGTCCACCGTCCTCCAGCCATGAGCCAGGCCTGTAACACAGAGCCCTGTC 1888

Qy       2111 CCCCCAGGTGGCATGTGGGCTCTTGGGGGCCCTGCTCAGCTACCTGTGGAGTTGGAATTC 2170
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1889 CCCCCAGGTGGCATGTGGGCTCTTGGGGGCCCTGCTCAGCTACCTGTGGAGTTGGAATTC 1948

Qy       2171 AGACCCGAGATGTGTACTGCCTGCACCCAGGGGAGACCCCTGCCCCTCCTGAGGAGTGCC 2230
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1949 AGACCCGAGATGTGTACTGCCTGCACCCAGGGGAGACCCCTGCCCCTCCTGAGGAGTGCC 2008

Qy       2231 GAGATGAAAAGCCCCATGCTTTACAAGCATGCAATCAGTTTGACTGCCCTCCTGGCTGGC 2290
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2009 GAGATGAAAAGCCCCATGCTTTACAAGCATGCAATCAGTTTGACTGCCCTCCTGGCTGGC 2068

Qy       2291 ACATTGAAGAATGGCAGCAGTGTTCCAGGACTTGTGGCGGGGGAACTCAGAACAGAAGAC 2350
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2069 ACATTGAAGAATGGCAGCAGTGTTCCAGGACTTGTGGCGGGGGAACTCAGAACAGAAGAC 2128

Qy       2351 TCACCTGTCGGCAGCTGCTAACGGATGGCAGCTTTTTGAATCTCTCAGATGAATTGTGCC 2410
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2129 TCACCTGTCGGCAGCTGCTAACGGATGGCAGCTTTTTGAATCTCTCAGATGAATTGTGCC 2188

Qy       2411 AAGGACCCAAGGCATCGTCTCACAAGTCCTGTGCCAGGACAGACTGTCCTCCACATTTAG 2470
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2189 AAGGACCCAAGGCATCGTCTCACAAGTCCTGTGCCAGGACAGACTGTCCTCCACATTTAG 2248

Qy       2471 CTGTGGGAGACTGGTCGAAGTGTTCTGTCAGTTGTGGTGTTGGAATCCAGAGAAGAAAGC 2530
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2249 CTGTGGGAGACTGGTCGAAGTGTTCTGTCAGTTGTGGTGTTGGAATCCAGAGAAGAAAGC 2308

Qy       2531 AGGTGTGTCAAAGGCTGGCAGCCAAAGGTCGGCGCATCCCCCTCAGTGAGATGATGTGCA 2590
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2309 AGGTGTGTCAAAGGCTGGCAGCCAAAGGTCGGCGCATCCCCCTCAGTGAGATGATGTGCA 2368

Qy       2591 GGGATCTACCAGGGTTCCCTCTTGTAAGATCTTGCCAGATGCCTGAGTGCAGTAAAATCA 2650
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2369 GGGATCTACCAGGGTTCCCTCTTGTAAGATCTTGCCAGATGCCTGAGTGCAGTAAAATCA 2428

Qy       2651 AATCAGAGATGAAGACAAAACTTGGTGAGCAGGGTCCGCAGATCCTCAGTGTCCAGAGAG 2710
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2429 AATCAGAGATGAAGACAAAACTTGGTGAGCAGGGTCCGCAGATCCTCAGTGTCCAGAGAG 2488

Qy       2711 TCTACATTCAGACAAGGGAAGAGAAGCGTATTAACCTGACCATTGGTAGCAGAGCCTATT 2770
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2489 TCTACATTCAGACAAGGGAAGAGAAGCGTATTAACCTGACCATTGGTAGCAGAGCCTATT 2548

Qy       2771 TGCTGCCCAACACATCCGTGATTATTAAGTGCCCCGTGCCGACGATTCCAGAAATCTCTGA 2830
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2549 TGCTGCCCAACACATCCGTGATTATTAAGTGCCCCGTGCCGACGATTCCAGAAATCTCTGA 2608
```

```
Qy       2831 TCCAGTGGGAGAAGGATGGCTTGGTTCTGCAGAATTCAAAGCTCTTGGATCACTAATT 2890
              | | ||||| ||||||||| |   |  |||||||||||||| || || | |||  |   |
Db       2609 TCCAGTGGGAGAAGGATGGTTTTTGGTGGAGAATTCAAAGGGATGGCATTACTCAATT 2668

Qy       2891 CAGGTTCACTAAAAATCCACGGTTCTTGCTGCTCCGACATCGGCGTGTACGGTGCATTG 2950
              |||||||||||||| |||||||| |||||||||||||||||||||||||||| | ||| |
Db       2669 CAGGTTCACTAAAAATTCACTGGTTTTGCTGGTCCCGACATCGGCGTGTACCGGTGCATTC 2728

Qy       2951 CAGGCTCTGCACAGGAAACACTTGTGCTCAAGCCATTGGTACTGACAACCGGCTCATCC 3010
              |||||||||||||||||| ||||||||||||| | ||||||||||||||| | ||  ||
Db       2729 CAGGCTCTGCACAGGAAACAGTTGTGCTCAAGCCATTGGTACTGACAACCGGCTCATCG 2788

Qy       3011 CACGCCCAGCCCTCAGGGAGCCTATGAGGGAATATCCTGGGATGGACCACAGCGAAGCCA 3070
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db       2789 CACGCCCAGCCCTCAGGGAGCCTATGAGGGAATATCCTGGGATGGACCACAGCGAAGCCA 2848

Qy       3071 ATAGTTTGGGAGTCACATGGCACAAAATGAGGCAAATGTGGAATAACAAAAATGACCTTT 3130
              |||||||||||||||||||||| |||||| ||||||||||||||||||||||| |||||
Db       2849 ATAGTTTGGGAGTCACATGGCACAAAATGAGGCAAATGTGGAATAACAAAAATGACCTTT 2908

Qy       3131 ATCTGGATGATGACCACATCAGTAACCAGCCTTCCTTGAGAGCTCTGTTAGGCCACTCCA 3190
              ||||||||||||||||||||||||||| |||||||||||||||||||||| |||| |||
Db       2909 ATCTGGATGATGACCACATCAGTAACCAGCCTTCCTTGAGAGCTCTGTTAGGCCACTTCA 2968

Qy       3191 GCAATTCTGCAGGAAGCACTAACTCCTGGGAGTTGAAGAATAAGCAGTTGGAAGCAGTAG 3250
              |||||||||||||||||| |||||||||||||||||||||||||||||||||||||| ||
Db       2969 GCAATTCTGCAGGAAGCACCAACTCCTGGGAGTTGAAGAATAAGCAGTTGGAAGCAGCAG 3028

Qy       3251 TTAAACAAGGAGCATATAGCATGGATACAGCCCAGTTTGATGAGCTGATAAGAAACATGA 3310
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3029 TTAAACAAGGAGCATATAGCATGGATACAGCCCAGTTTGATGAGCTGATAAGAAACATGA 3088

Qy       3311 GTCAGCTCATGGAAACCGGAGAGGTCAGCGATGATCTTGCGTCCCAGCTGATATATCAGC 3370
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3089 GTCAGCTCATGGAAACCGGAGAGGTCAGCGATGATCTTGCGTCCCAGCTGATATATCAGC 3148

Qy       3371 TGGTGGCCGAATTAGCCAAGGCACAGCCAACACACATGCAGTGGCGGGGCATCCAGGAAG 3430
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||   |
Db       3149 TGGTGGCCGAATTAGCCAAGGCACAGCCAACACACATGCAGTGGCGGGGCATCCAGGAAG 3208

Qy       3431 AGACACCTCCTGCTGCTCAGCTCAGAGGGGAAACAGGGAGTGTGTCCCAAAGCTGGCATG 3490
              ||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db       3209 AGACACCTCCTGCTGCTCAGCTCAGAGGGGAAACAGGGAGTGTGTCCCAAAGCTGGCATG 3268

Qy       3491 CAAAAACTCAGGCAAGCTGACATTCAAGCCGAAAGGACCTGTTCTCATGAGGCAAAGCT 3550
              |||||||||||||||||   ||||||||||||||||||||||||||||||||||||||
Db       3269 CAAAAACTCAGGCAAGCTGACATTCAAGCCGAAAGGACCTGTTCTCATGAGGCAAAGCT 3328

Qy       3551 AACCTCCCTCAATTTCATTTAATAAAACAATAAATTCCAGGATTGGAAATACAGTATACA 3610
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db       3329 AACCTCCCTCAATTTCATTTAATAAAACAATAAATTCCAGGATTGGAAATACAGTATACA 3388

Qy       3611 TTACAAAAAGGACAGAGGTCATCAATATACTGTGTGACCTTATTACCCCAGTGAGGCCA 3670
              ||||||||||||||||||||||||||||||||||||||||||| | | | | | | |
Db       3389 TTACAAAAAGGACAGAGGTCATCAATATACTGTGTGACCTTATTACCCCAGTGAGGCCA 3448

Qy       3671 CATATACATGGACTAAGGATGGAACCTTGTTACAGCCTCAGTAAAAATAATTTTGGATG 3730
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| || |
Db       3449 CATATACATGGACTAAGGATGGAACCTTGTTACAGCCTCAGTAAAAATAATTTTGGATG 3508

Qy       3731 GAACTGGGAAGATACAGATACAGAATCCTACAAGGAAAGAACAAGGCATATATGAATCTT 3790
              || |||||||||||||||| |||||||||||||||||||||||||||| |  || | |
Db       3509 GAACTGGGAAGATACAGATACAGAATCCTACAAGGAAAGAACAAGGCATATATGAATCTT 3568

Qy       3791 CTGTAGCTAATCATCTTGGTTCAGATGTGGAAAGTTCTTCTGTGCTGTATGCAGAGGCAC 3850
              |||||||||||||||||||||||||||||||||||||||||||||||||| || |  |
Db       3569 CTGTAGCTAATCATCTTGGTTCAGATGTGGAAAGTTCTTCTGTGCTGTATGCAGAGGCAC 3628

Qy       3851 CTGTTATCTTGTCTGTTCAAAGAAATATCACCAAACCAGAGCACAACCACTCCGTCTTGC 3910
              ||||||||||| |  | | |  |||||||||||||||||||||||| ||| | |
Db       3629 CTGTTATCTTGTTCTGTTGAAAGAAATATCACCAAACCAGAGCACAACCACTCCGTCTTGC 3688

Qy       3911 TGGTTGGAGGCATGGTGGAGCCACCTTGGACCAAAGGGACAATCCATGGCCTGTAA 3970
              ||||||||||||||| |||||   |  |  |  |   ||| ||||||||||||
Db       3689 TGGTTGGAGGCATGGTGGAGCCACCTTGGACCAAAGGGACAATCCATGGCCTGTAA 3748

Qy       3971 AAGGTGTCCTCAGCCTAATATAACTTGGTTGAAGAGAGGAGGATCCTGAGTGGCAATG 4030
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3749 AAGGTGTCCTCAGCCTAATATAACTTGGTTGAAGAGAGGAGGATCCTGAGTGGCAATG 3808

Qy       4031 TTTCTTGCTTTTCAATGGATCCTGTTTGTTGCAGAATCTTTCTTTGAAAATGAAGGAA 4090
              |||||||||| ||||||| |  ||||||||| |  | ||||| |||||||||||||
Db       3809 TTTCTTGCTTTTCAATGGATCCTGTTTGTTGCAGAATCTTTCTTTGAAAATGAAGGAA 3868

Qy       4091 CCTACGTCTGCATAGCCACCAATGCTCTTGGAAAGGCAGTGGCAACATCTGTACTCCACT 4150
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3869 CCTACGTCTGCATAGCCACCAATGCTCTTGGAAAGGCAGTGGCAACATCTGTACTCCACT 3928

Qy       4151 TGCTGGAACGAAGATGGCCAGAGAGTAGAATCGTATTTCTGGAAGGACATAAAAAGTACA 4210
              |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db       3929 TGCTGGAACGAAGATGGCCAGAGAGTAGAATCGTATTTCTGGAAGGACATAAAAAGTACA 3988

Qy       4211 TTCTCCAGGCAACCAACACTAGAACCAACAGCAATGACCCAACAGGAGAACCCCCGCCTC 4270
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3989 TTCTCCAGGCAACCAACACTAGAACCAACAGCAATGACCCAACAGGAGAACCCCCGCCTC 4048

Qy       4271 AAGAGGCTTTTTGGGAGCCTGGTAACTGGTACATTGGTCTGGCAGCTGGTGGTCATTGG 4330
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4049 AAGAGGCTTTTTGGGAGCCTGGTAACTGGTACATTGGTCTGGCAGCTGGTGGTCATTGG 4108

Qy       4331 GAGCCCGGCATTCAGAGACCCCAGTGTGTGATGGCCAATGGGCAGGAAGTGAGTGAGGCCC 4390
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4109 GAGCCCGGCATTCAGAGACCCCAGTGTGTGATGGCCAATGGGCAGGAAGTGAGTGAGGCCC 4168

Qy       4391 TGTGTGATCACCTCCAGAAGCACTGGCTGGGTTGAGCCCTGTAACATCCGGGACTGCC 4450
              |||||||||||||||||||||||||||||||||||||  | | |  ||||||||||||
Db       4169 TGTGTGATCACCTCCAGAAGCACTGGCTGGGTTGAGCCCTGTAACATCCGGGACTGCC 4228

Qy       4451 CAGCCGAGGTGGTTCACAAGTGTGTGGTCACAGTGCTCTGTGTCTTGCGGTGAAGGATACC 4510
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4229 CAGCCGAGGTGGTTCACAAGTGTGTGGTCACAGTGCTCTGTGTCTTGCGGTGAAGGATACC 4288

Qy       4511 ACAGTCGGCAGGTGACGTGCAAGCGGACAAAAGCCAATGGAACTGTGCAGGTGGTGTCTC 4570
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4289 ACAGTCGGCAGGTGACGTGCAAGCGGACAAAAGCCAATGGAACTGTGCAGGTGGTGTCTC 4348

Qy       4571 CAAGAGCATGTGCCCCTAAAGACCGGCCTCTGGGAAGAAAACCATGTTTTGGTCATCCAT 4630
              |||||||||||||||||||||||||||||||||||||||||  ||||||||||||||||||
Db       4349 CAAGAGCATGTGCCCCTAAAGACCGGCCTCTGGGAAGAAAACCATGTTTTGGTCATCCAT 4408

Qy       4631 GTGTTCAGTGGGAACCAGGGAACCGGTGTCCTGGACGTTGCATGGGCCGTGCTGTGAGGA 4690
              |||||||||||  ||||||||||||||||||||||||||||||  ||||||||||||||||
Db       4409 GTGTTCAGTGGGAACCAGGGAACCGGTGTCCTGGACGTTGCATGGGCCGTGCTGTGAGGA 4468

Qy       4691 TGCAGCAGCGTCACACAGCTTGTCAACACAACAGCTCTGACTCCAACTGTGATGACAGAA 4750
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4469 TGCAGCAGCGTCACACAGCTTGTCAACACAACAGCTCTGACTCCAACTGTGATGACAGAA 4528

Qy       4751 AGAGACCCACCTTAAGAAGGAACTGCACATCAGGGGCCTGGATGTGGTTGGCACACAG 4810
              |||||||||||||||||||||||| ||||||||| ||| || |||||||||||||||||
Db       4529 AGAGACCCACCTTAAGAAGGAACTGCACATCAGGGGCCTGGATGTGGTTGGCACACAG 4588

Qy       4811 GCCCTTGGAAGCCCTGTACAGCAGCCTGTGGCAGGGGTTCCAGTCTTGGAAAGTCGACT 4870
              ||||||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db       4589 GCCCTTGGAAGCCCTGTACAGCAGCCTGTGGCAGGGGTTCCAGTCTTGGAAAGTCGACT 4648

Qy       4871 GTATCCACACAAGGAGTTGCAAACCTGTGGCCAAGAGACACTTCCTACAGAAAAAGAAAC 4930
              ||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db       4649 GTATCCACACAAGGAGTTGCAAACCTGTGGCCAAGAGACACTTCCTACAGAAAAAGAAAC 4708

Qy       4931 CAATTTCCTGGCGGCACTGTCTTGGGCCTCCTGTGATAGAGACTGCACAGACACAACTC 4990
              ||||||||||||||||||||||||||||||||||||||| || ||||| | ||||||
Db       4709 CAATTTCCTGGCGGCACTGTCTTGGGCCTCCTGTGATAGAGACTGCACAGACACAACTC 4768

Qy       4991 ACTACTGTACCCTGTAAAACATCCTAATTTCCCTTTCTCAGACTTCCTACAAATAAAGTT 5050
```

```
                ||||||||||||||||:||||:||||||||||||||||||||||||||||||||||||||
Db       4769 ACTACTGTATGTTTGTAAAACATCTTAATTTCTCTTCTCTACACCCTACAAACAAAGGT 4828

Qy       5051 GCTGCCAGTCATGTCAAGAGGGATAA 5076
              ||||||||||||||||||||||||||
Db       4829 GCTGCCAGTCATGTCAAGAGGGATAA 4854


RESULT 4
AAD17758
ID   AAD17758 standard; DNA; 6303 BP.
XX
AC   AAD17758;
XX
DT   10-DEC-2001  (first entry)
XX
DE   Human novel KIAA1233-like protein, NOV-2b encoding DNA.
XX
KW   Human; NOV-X protein; KIAA1233-like protein; STE20-like protein; tumour;
KW   trypsin inhibitor-like protein; gene therapy; haematopoietic; illness;
KW   immunological disorder, neurodegenerative disorder; Alzheimer's disease;
KW   Parkinson's disease; immunomodulatory, pharmacogenomic, haemostatic;
KW   human immunodeficiency virus; HIV; fertility disorder, neuroprotective,
KW   cytostatic; nootropic; anti-infertility; cancer; chromosome 15;
KW   NOV-2b protein; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   5'UTR           1..424
FT                   /*tag=  a
FT   CDS             425..4270
FT                   /*tag=  b
FT                   /product= "Human novel KIAA1233 like protein, NOV-2b"
FT   sig_peptide     425..502
FT                   /*tag=  c
FT   mat_peptide     503..4267
FT                   /*tag=  d
FT                   /product= "Human mature novel KIAA1233-like protein,
FT                   NOV-2b"
FT   3'UTR           4271..6303
FT                   /*tag=  e
XX
PN   WO200162928-A2.
XX
PD   30-AUG-2001.
XX
PF   26-FEB-2001; 2001WO-US06151.
XX
PR   25-FEB-2000; 2000US-0184951.
PR   28-FEB-2000; 2000US-0185548.
PR   01 MAR-2000; 2000US-0185967.
PR   18-APR-2000; 2000US-0197723.
PR   27-APR-2000; 2000US-0199957.
PR   23-FEB-2001; 2001US-0789390.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Vernet CAM,  Fernandes E,  Shimkets RA,  Macdougall J,  Spaderna SK;
XX
DR   WPI; 2001-582051/65.
DR   P-PSDB; AAE10610.
XX
PT   New isolated KIAA1233-like, STE20-like, or trypsin inhibitor-like
PT   polypeptide for diagnosing and treating pathological disorders, such as
PT   Parkinson's disease and for use in pharmacogenomics -
XX
PS   Claim 9; Page 35-37; 189pp; English.
XX
CC   The invention relates to novel human polypeptides referred as NOV-X
CC   and their corresponding nucleic acid sequences. NOV X collectively
CC   include NOV-1, NOV-2a and NOV 2b which are novel KIAA1233 like
CC   polypeptides, NOV-3a, NOV 3b, NOV 3c and NOV-3d which are novel
```

```
CC   STE20 like polypeptides and NOV 4a, NOV 4b, NOV 4c, NOV-4d and NOV-4e
CC   which are novel trypsin inhibitor like polypeptides. NOV-X is used to
CC   identify a potential therapeutic agent that can modulate its activity
CC   and can be used for treating a pathology related to aberrant expression
CC   or aberrant physiological interactions of NOV X. NOV X or its DNA is
CC   used to determine the presence or predisposition to a disease associated
CC   with altered levels of NOV-X. NOV-X, its DNA and its antibody are used
CC   to treat or prevent a pathology associated with NOV X. The pathological
CC   states that can be treated or prevented are haematopoietic, cancer,
CC   immunological, tumour, neurodegenerative (e.g. Alzheimer's and
CC   Parkinson's disease), human immunodeficiency virus (HIV) illness and
CC   fertility disorders. NOV X and its DNA are used in pharmacogenomics for
CC   predictive medicine. NOV X DNA is used in gene therapy. The present
CC   sequence is a DNA encoding human novel KIAA1233 like protein, NOV-2b
CC   NOV-2b gene is located at chromosome 15.
XX
SQ   Sequence 6303 BP; 1792 A; 1297 C; 1528 G; 1686 T; 0 other;

  Query Match             76.1%,  Score 3862.2,  DB 22;  Length 6303;
  Best Local Similarity   99.9%,  Pred. No. 0;
  Matches 3864;  Conservative    0,  Mismatches    3,  Indels    0,  Gaps    0;

Qy       1210 AGTGATGGATTTAAAGAGATAATGCCCTATGACCACTTCCAACCTCTTCCTCGCTGGGAA 1269
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        404 AGTGATGGATTTAAAGAGATAATGCCCTATGACCACTTCCAACCTCTTCCTCGCTGGGAA 463

Qy       1270 CATAATCCTTGGACTGCATGTTCCGTGTCCTGTGGAGGAGGATTCAGAGACGGAGCTTT 1329
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        464 CATAATCCTTGGACTGCATGTTCCGTGTCCTGTGGAGGAGGATTCAGAGACGGAGCTTT 523

Qy       1330 GTGTGTGTAGAGGAATCCATGCATGGAGAGATATTGCAGGTGGAAGAATGGAAGTGCATG 1389
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        524 GTGTGTGTAGAGGAATCCATGCATGGAGAGATATTGCAGGTGGAAGAATGGAAGTGCATG 583

Qy       1390 TACGCACCCAAACCCAAGGTTATGCAAACTTGTAATCTGTTTGATTGCCCCAAGTGGATT 1449
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        584 TACGCACCCAAACCCAAGGTTATGCAAACTTGTAATCTGTTTGATTGCCCCAAGTGGATT 643

Qy       1450 GCCATGGAGTGGTCTCAGTGCACAGTGACTTGTGGCCGAGGGTTACGGTACCGGGTTGTT 1509
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        644 GCCATGGAGTGGTCTCAGTGCACAGTGACTTGTGGCCGAGGGTTACGGTACCGGGTTGTT 703

Qy       1510 CTGTGTATTAACCACCGCGGAGAGCATGTTGGGGGCTGCAATCCACAACTGAAGTTACAC 1569
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        704 CTGTGTATTAACCACCGCGGAGAGCATGTTGGGGGCTGCAATCCACAACTGAAGTTACAC 763

Qy       1570 ATCAAAGAAGAATGTGTCATTCCCATCCCGTGTTATAAACCAAAAGAAAAAAGTCCAGTG 1629
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        764 ATCAAAGAAGAATGTGTCATTCCCATCCCGTGTTATAAACCAAAAGAAAAAAGTCCAGTG 823

Qy       1630 GAAGCAAAATTGCCTTGGCTGAAACAAGCACAAGAACTAGAAGAGACCAGAATAGCAACA 1689
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        824 GAAGCAAAATTGCCTTGGCTGAAACAAGCACAAGAACTAGAAGAGACCAGAATAGCAACA 883

Qy       1690 GAAGAACCAACGTTCATTCCAGAACCCTGGTCAGCCTGCAGTACCACGTGTGGGCCAGGT 1749
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db        884 GAAGAACCAACGTTCATTCCAGAACCCTGGTCAGCCTGCAGTACCACGTGTGGGCCGGGT 943

Qy       1750 GTGCAGGTCCGCGAGGTGAAGTGCCGTGTGCTCCTCACATTCACGCAGACTGAGACTGAG 1809
              ||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        944 GTGCAGGTCCGTGAGGTGAAGTGCCGTGTGCTCCTCACATTCACGCAGACTGAGACTGAG 1003

Qy       1810 CTGCCCGAGGAAGAGTGTGAAGGCCCCAAGCTGCCCACCGAACGGCCCTGCCTCCTGGAA 1869
              ||||||||||||||||||||||||||.|||||||||||||||||||||||||||||||||
Db       1004 CTGCCCGAGGAAGAGTGTGAAGGCCCCAAGCTGCCCACCGAACGGCCCTGCCTCCTGGAA 1063

Qy       1870 GCATGTGATGAGAGCCCGGCCTCCCGAGAGCTAGACATCCCTCTCCCTGAGGACAGTGAG 1929
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1064 GCATGTGATGAGAGCCCGGCCTCCCGAGAGCTAGACATCCCTCTCCCTGAGGACAGTGAG 1123

Qy       1930 ACGACTTACGACTGGCAGTACGCTGGGTTCACCCCTTGCACAGCAACATGCTTGGGAGGC 1989
              ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db       1124 ACGACTTACGACTGGCAGTACGCTGGGTTCACCCCTTGCACAGCAACATGCCTGGGAGGC 1183
```

```
Qy       1990 CATCAAGAAGGTATAGTAGTCGGTTTACATATCCAGACCAGGAGACAGTCAATGACAGCC 2049
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1184 CATCAAGAAGGTATAGTAGTTGGTTTACATATCCAGACCAGGAGACAGTCAATGACAGCC 1243

Qy       2050 TTGTGTGATATGGTTCACCTTCCTCCAGCCATGAGCCAGGCCTGTAACACAGAGGCCTGT 2109
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1244 TTGTGTGATATGGTCCACCGTCCTCCAGCCATGAGCCAGGCCTGTAACACAGAGGCCTGT 1303

Qy       2110 CCCCCCAGGTGGCATGTGGGCTCTTGGGGGCCCTGCTCAGCTACCTGTGGAGTTGGAATT 2169
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1304 CCCCCCAGGTGGCATGTGGGCTCTTGGGGGCCCTGCTCAGCTACCTGTGGAGTTGGAATT 1363

Qy       2170 CAGACCCGAGATGTGTACTGCCTGCACCCAGGGGAGACCCCTGCCCCTCCTGAGGAGTGC 2229
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1364 CAGACCCGAGATGTGTACTGCCTGCACCCAGGGGAGACCCCTGCCCCTCCTGAGGAGTGC 1423

Qy       2230 CGAGATGAAAAGCCCCATGGTTTACAAGCATGCAATCAGTTTGACTGGCCTCCTGGCTGG 2289
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1424 CGAGATGAAAAGCCCCATGGTTTACAAGCATGCAATCAGTTTGACTGCCCCTCCTGGCTGG 1483

Qy       2290 CACATTGAAGAATGGCAGCAGTGCTCCAGGACTTGTGGCGGGGGAACTCAGAACAGAAGA 2349
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1484 CACATTGAAGAATGGCAGCAGTGCTCCAGGACTTGTGGCGGGGGAACTCAGAACAGAAGA 1543

Qy       2350 GTCACCTGTCGGTAGCTGCTAATGGATGGTAGCTTTTTGAATCTCTCAGATGAATTGTCC 2409
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1544 GTCACCTGTCGGTAGCTGCTAATGGATGGTAGCTTTTTGAATCTCTCAGATGAATTGTCC 1603

Qy       2410 CAAGGACCCAAGGTATGCCCTACAAGTCTGTGCCAGGACAGACTGGCTTCCCCACATTTA 2469
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1604 CAAGGACCCAAGGTATGCCCTACAAGTCTGTGCCAGGACAGACTGGCTTCCCCACATTTA 1663

Qy       2470 GCTGTGGGAGACTGGTGGAAGTGCTCTGTCAGTTGTGGTGTTGGAATCCAGAGAAGAAAT 2529
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1664 GCTGTGGGAGACTGGTGGAAGTGCTCTGTCAGTTGTGGTGTTGGAATCCAGAGAAGAAAT 1723

Qy       2530 CAGGTGTGTCAAAGGCTGGCAGCCAAAGGTGGGCGCATCCCCCTCAGTGAGATGATGTGC 2589
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1724 CAGGTGTGTCAAAGGCTGGCAGCCAAAGGTGGGCGCATCCCCCTCAGTGAGATGATGTGC 1783

Qy       2590 AGGGATCTACCAGGGTTCCCTCTTGTAAGATCTTGCCAGATGCCTGAGTGCAGTAAAATC 2649
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1784 AGGGATCTACCAGGGTTCCCTCTTGTAAGATCTTGCCAGATGCCTGAGTGCAGTAAAATC 1843

Qy       2650 AAATCAGAGATGAAGACAAAACTTGGTGAGCAGGGTCCGCAGATCCTCAGTGTCCAGAGA 2709
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1844 AAATCAGAGATGAAGACAAAACTTGGTGAGCAGGGTCCGCAGATCCTCAGTGTCCAGAGA 1903

Qy       2710 GTCTACATTCAGACAAGGGAAGAGAAGCGTATTAACCTGACCATTGGTAGCAGAGCCTAT 2769
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1904 GTCTACATTCAGACAAGGGAAGAGAAGCGTATTAACCTGACCATTGGTAGCAGAGCCTAT 1963

Qy       2770 TTGCTGCCCAACACATCCGTGATTATTAAGTGCCCCGTGCGACGATTCCAGAAATCTCTC 2829
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1964 TTGCTGCCCAACACATCCGTGATTATTAAGTGCCCCGTGCGACGATTCCAGAAATCTCTC 2023

Qy       2830 ATCCATGTGGAGAAGGATGGCCCCTGGCCTCCAGAACTCCAAAGGCCTTGCATCACCAAT 2889
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2024 ATCCATGTGGAGAAGGATGGCCCCTGGCCTCCAGAACTCCAAAGGCCTTGCATCACCAAT 2083

Qy       2890 TCAGGCTCACTAAAAATCCACTTGCTGGCTGCCCCCGACATGCGGTGGTACCGGTGCATT 2949
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2084 TCAGGCTCACTAAAAATCCACTTGCTGGCTGCCCCCGACATGCGGTGGTACCGGTGCATT 2143

Qy       2950 GCAGGCTCTGCACAGGAAACAGCCAGTGCTAAGCTCATTGGTACTGACAACCGGCTCATC 3009
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2144 GCAGGCTCTGCACAGGAAACAGCCAGTGCTAAGCTCATTGGTACTGACAACCGGCTCATC 2203

Qy       3010 GCACGCCCAGCCCTCAGGGAGCCTATGAGGGAATATCCTGGGATGGACCACAGCGAAGCC 3069
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2204 GCACGCCCAGCCCTCAGGGAGCCTATGAGGGAATATCCTGGGATGGACCACAGCGAAGCC 2263

Qy       3070 AATAGTTTGGGAGTCACATGGTACAAAATGAGGTAAATGTGGAATAACAAAAATGACCTT 3129
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2264 AATAGTTTGGGAGTCACATGGTACAAAATGAGGTAAATGTGGAATAACAAAAATGACCTT 2323

Qy       3130 TATCTGGATGATGACCACATTAGTAACCAGCCTTTCTTGAGAGCTCTGTTAGGCCACTGC 3189
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2324 TATCTGGATGATGACCACATTAGTAACCAGCCTTTCTTGAGAGCTCTGTTAGGCCACTGC 2383

Qy       3190 AGCAATTCTGCAGGAAGCACCAACTCCTGGGAGTTGAAGAATAAGCAGTTTGAAGCAGCA 3249
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2384 AGCAATTCTGCAGGAAGCACCAACTCCTGGGAGTTGAAGAATAAGCAGTTTGAAGCAGCA 2443

Qy       3250 GTTAAACAAGGAGCATATAGCATGGATACAGCCCAGTTTGATGAGCTGATAAGAAACATG 3309
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2444 GTTAAACAAGGAGCATATAGCATGGATACAGCCCAGTTTGATGAGCTGATAAGAAACATG 2503

Qy       3310 AGTCAGCTCATGGAAACGGAGAGGTCAGCGATGATCTTGGCGTCCCAGCTGATATATCAG 3369
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2504 AGTCAGCTCATGGAAACGGAGAGGTCAGCGATGATCTTGGCGTCCCAGCTCATATATCAG 2563

Qy       3370 CTGGTGGCCGAATTAGCCAAGGCACAGCCAACAGACATCCAGCGGGGGGGCATCCAGGAA 3429
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2564 CTGGTGGCCGAATTAGCCAAGGCACAGCCAACAGACATCCAGCGGGGGGGCATCCAGGAA 2623

Qy       3430 GAGACACCTCCTGCTGCCTCAGCTCAGAGGGGAAACAGGGAGTCCGTCCAAAGCTCGCAT 3489
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2624 GAGACACCTCCTGCTGCCTCAGCTCAGAGGGGAAACAGGGAGTCCGTCCAAAGCTCGCAT 2683

Qy       3490 GCAAAAAACTCAGGCAAGCTGACATTCAAGCCGAAAGGACCTGCTTCTCATGACCAAAGC 3549
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2684 GCAAAAAACTCAGGCAAGCTGACATTCAAGCCGAAAGGACCTGCTTCTCATGACCAAAGC 2743

Qy       3550 CAACCTCCCTCAATTTCATTTAATAAAACAATAAATTCCAGGATTGGAAATACAGTATAC 3609
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2744 CAACCTCCCTCAATTTCATTTAATAAAACAATAAATTCCAGGATTGGAAATACAGTATAC 2803

Qy       3610 ATTACAAAAAGGACAGAGGTCATCAATATACTGTGTGACCTTATTACCCCCAGTGAGGCC 3669
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2804 ATTACAAAAAGGACAGAGGTCATCAATATACTGTGTGACCTTATTACCCCCAGTGAGGCC 2863

Qy       3670 ACATATACATGGACCAAGGATGGAACCTTGTTACAGCCCTCAGTAAAAATAATTTTGGAT 3729
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2864 ACATATACATGGACCAAGGATGGAACCTTGTTACAGCCCTCAGTAAAAATAATTTTGGAT 2923

Qy       3730 GGAACTGGCAAGATACAGATACAGAATCCTACAAGGAAAGAACAAGGCATATATGAATGT 3789
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2924 GGAACTGGCAAGATACAGATACAGAATCCTACAAGGAAAGAACAAGGCATATATGAATGT 2983

Qy       3790 TCTGTAGCTAATCATCTTGGTTCAGATGTGGAAAGTTCTTCTTGTCTCCATGCAGAGGCA 3849
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2984 TCTGTAGCTAATCATCTTGGTTCAGATGTGGAAAGTTCTTCTTGTCTCCATGCAGAGGCA 3043

Qy       3850 CCTGTCATCTTGTCTGTTGAAACAAAATATCACCAAACTACACACAACCATCTGTCTGTT 3909
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3044 CCTGTCATCTTGTCTGTTGAAACAAAATATCACCAAACTACACACAACCATCTGTCTGTT 3103

Qy       3910 GTGGTTGGAGGCATCGTGGAGGCAGCCCTTGGTAGCAAACCTCACAATCCGATGCCTGTA 3969
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3104 GTGGTTGGAGGCATCGTGGAGGCAGCCCTTGGTAGCAAACCTCACAATCCGATGCCTGTA 3163

Qy       3970 AAAGGTGTCCCTCAGCCTAATATAACTTGGTTGAAGAGAGGATGATCCTCCAGGGGCAAT 4029
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3164 AAAGGTGTCCCTCAGCCTAATATAACTTGGTTGAAGAGAGGAGGATCTCTGAGGGGCAAT 3223

Qy       4030 GTTTCCTTGCTTTTCAATGGATCCCTGTTGTTGCAGAATGTTTCCCTTGAAAATGAAGGA 4089
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3224 GTTTCCTTGCTTTTCAATGGATCCCTGTTGTTGCAGAATGTTTCCCTTGAAAATGAAGGA 3283

Qy       4090 ACCTACGTCTGCATAGCCACCAATGCTCTTGGAAAGGCAGTGGCAACATCTGTACTCCAC 4149
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3284 ACCTACGTCTGCATAGCCACCAATGCTCTTGGAAAGGCAGTGGCAACATCTGTACTCCAC 3343

Qy       4150 TTGCTGGAACGAAGATGGCCAGACAGATAGAATCTATCTTTGAAGGAGACAAAAGAGTAC 4209
```

```
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3344 TTGCTGGAACGAAGATGGCCAGAGAGTAGAATCGTATTTCTGCAAGGACATAAAAAGTAC 3403

Qy       4210 ATTCTCCAGGCAACCAACACTAGAACCAACAGCAATGACCCAACAGGAGAACCCCGCCT 4269
               |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db       3404 ATTCTCCAGGCAACCAACACTAGAACCAACAGCAATGACCCAACAGGACAACCCCGCCT 3463

Qy       4270 CAAGAGCCTTTTTGGGAGCCTGGTAACTGGTCACATTGTTCTGCCACCTGTGGTCATTTG 4329
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3464 CAAGAGCCTTTTTGGGAGCCTGGTAACTGGTCACATTCTTCTGCCACCTGTGGTCATTTG 3523

Qy       4330 GGAGCCGGCATTCAGAGACCCCAGTGTGTGATGGCCAATGGGCAGGAAGTGAGTGAGGCC 4389
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3524 GGAGCCGGCATTCAGAGACCCCAGTGTGTGATGGCCAATGGGCAGGAAGTGAGTGAGGCC 3583

Qy       4390 CTGTGTGATCACCTCCAGAAGCCACTGGCTGGCTTTGAGCCCTGTAACATCCGGGACTGC 4449
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3584 CTGTGTGATCACCTCCAGAAGCCACTGGCTGGGTTTGAGCCCTGTAACATCCGGGACTGC 3643

Qy       4450 CCAGCGAGGTGGTTCACAAGTGTGTGGTCACAGTGCTCTGTGTCTTGCGGTGAAGGATAC 4509
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3644 CCAGCGAGGTGGTTCACAAGTGTGTGGTCACAGTGCTCTGTGTCTTGCGGTGAAGGATAC 3703

Qy       4510 CACAGTCGGCAGGTGACGTGCAAGCGGACAAAAGCCAATGGAACTGTGCAGGTGGTGTCT 4569
               |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db       3704 CACAGTCGGCAGGTGACGTGCAAGCGGACAAAAGCCAATGGAACTGTGCAGGTGGTGTCT 3763

Qy       4570 CCAAGAGCATGTGCCCCTAAAGACCGGCCTCTGGAAGAAAACCATGTTTTGGTCATCCA 4629
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3764 CCAAGAGCATGTGCCCCTAAAGACCGGCCTCTGGAAGAAAACCATGTTTTGGTCATCCA 3823

Qy       4630 TGTGTTCAGTGGGAACCAGGGAACCGGTGTCCTGGACCTTCCATGGCCCGTGCTGTGAGC 4689
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3824 TGTGTTCAGTGGGAACCAGGGAACCGGTGTCCTGGACGTTGCATGGGCCGTGCTGTGAGG 3883

Qy       4690 ATGCAGCAGCGTCACACAGCTTGTCAACACAACAGCTCTGACTCCAACTGTGATGACAGA 4749
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3884 ATGCAGCAGCGTCACACAGCTTGTCAACACAACAGCTCTGACTCCAACTGTGATGACAGA 3943

Qy       4750 AAGAGACCCACCTTAAGAAGGAACTGCACATCAGGGGCCTGTGATGTGTGTTGGCACACA 4809
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3944 AAGAGACCCACCTTAAGAAGGAACTGCACATCAGGGGCCTGTGATGTGTGTTGGCACACA 4003

Qy       4810 GGCCCTTGGAAGCCCTGTACAGCAGCCTGTGGCAGGGGTTTCCAGTCTCGGAAAGTCGAC 4869
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4004 GGCCCTTGGAAGCCCTGTACAGCAGCCTGTGGCAGGGGTTTCCAGTCTCGGAAAGTCGAC 4063

Qy       4870 TGTATCCACACAACGACTTGCAAACCTGTGGCCAACACACACTGTGTACAGAAAAAGAAA 4929
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4064 TGTATCCACACAACGACTTGCAAACCTCTGGCCAAGAGACACTGTGTACAGAAAAACAAA 4123

Qy       4930 CCAATTTCCTGGCGGCACTGTCTTGGGCCCTCCTGTGATAGAGACTGCACAGACACAACT 4989
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4124 CCAATTTCCTGGCGGCACTGTCTTGGGCCCTCCTGTGATAGAGACTGCACAGACACAACT 4183

Qy       4990 CACTACTGTATGTTTGTAAAACATCTTAATTTGTGTTCTCTAGACCGCTACAAACAAAGG 5049
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4184 CACTACTGTATGTTTGTAAAACATCTTAATTTCTCTTCTCTAGACCCCTACAAACAAACC 4243

Qy       5050 TGCTGCCAGTCATGTCAAGAGGGATAA 5076
               |||||||||||||||||||||||||||
Db       4244 TGCTGCCACTCATCTCAAGAGGGATAA 4270


RESULT 5
AAD17756
ID   AAD17756 standard; DNA; 6294 BP.
XX
AC   AAD17756;
XX
DT   10-DEC-2001  (first entry)
XX
DE   Human novel KIAA1233 like protein, NOV 1 encoding DNA.
XX
KW   Human; NOV-X protein; KIAA1233-like protein; STE20-like protein; tumour;
KW   trypsin inhibitor like protein, gene therapy, haematopoietic, illness;
KW   immunological disorder; neurodegenerative disorder; Alzheimer's disease;
KW   Parkinson's disease, immunomodulatory, pharmacogenomic; haemostatic;
KW   human immunodeficiency virus; HIV; fertility disorder; neuroprotective;
KW   cytostatic; nootropic; anti-infertility; cancer; chromosome 15;
KW   NOV-1 protein; ds
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   5'UTR           1..415
FT                   /*tag=  a
FT   CDS             416..4261
FT                   /*tag=  b
FT                   /product= "Human novel KIAA1233 like protein, NOV-1"
FT   3'UTR           4262..6294
FT                   /*tag=  c
XX
PN   WO200162928-A2.
XX
PD   30-AUG-2001.
XX
PF   26-FEB-2001; 2001WO-US06151.
XX
PR   25-FEB-2000; 2000US-0184951.
PR   28-FEB-2000; 2000US-0185548.
PR   01-MAR-2000; 2000US-0185967.
PR   18-APR-2000; 2000US-0197723.
PR   27-APR-2000; 2000US-0199957.
PR   23-FEB-2001; 2001US-0789390.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Vernet CAM,  Fernandes E,  Shimkets RA,  Macdougall J,  Spaderna SK;
XX
DR   WPI; 2001-582051/65.
DR   P-PSDB; AAE10608.
XX
PT   New isolated KIAA1233 like, STE20-like, or trypsin inhibitor-like
PT   polypeptide for diagnosing and treating pathological disorders, such as
PT   Parkinson's disease and for use in pharmacogenomics -
XX
PS   Claim 9; Page 4-7; 189pp; English.
XX
CC   The invention relates to novel human polypeptides referred as NOV X
CC   and their corresponding nucleic acid sequences. NOV X collectively
CC   include NOV-1, NOV-2a and NOV-2b which are novel KIAA1233-like
CC   polypeptides, NOV 3a, NOV 3b, NOV 3c and NOV 3d which are novel
CC   STE20-like polypeptides and NOV 4a, NOV-4b, NOV-4c, NOV 4d and NOV-4e
CC   which are novel trypsin inhibitor like polypeptides. NOV X is used to
CC   identify a potential therapeutic agent that can modulate its activity
CC   and can be used for treating a pathology related to aberrant expression
CC   or aberrant physiological interactions of NOV X. NOV X or its DNA is
CC   used to determine the presence or predisposition to a disease associated
CC   with altered levels of NOV X. NOV X, its DNA and its antibody are used
CC   to treat or prevent a pathology associated with NOV-X. The pathological
CC   states that can be treated or prevented are haematopoietic, cancer,
CC   immunological, tumour, neurodegenerative (e.g. Alzheimer's and
CC   Parkinson's disease), human immunodeficiency virus (HIV) illness and
CC   fertility disorders. NOV X and its DNA are used in pharmacogenomics for
CC   predictive medicine. NOV-X DNA is used in gene therapy. The present
CC   sequence is a DNA encoding human novel KIAA1233-like protein, NOV 1.
CC   NOV-1 gene is located at chromosome 15.
XX
SQ   Sequence 6294 BP; 1789 A; 1297 C; 1528 G; 1680 T; 0 other;

  Query Match             76.0%;  Score 3859;  DB 22;  Length 6294;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 3862;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1210 AGTGATGGATTTAAAGAGATAATGGCCATGACCATTCCAACCTCTTCCTCGCTGGGAA 1269
```

```
              ||||||||||||||||||||   |||||  ||||||||||||||||| |||||| ||||||
Db        395 AGTGATGGATTTAAAGAGATAATGCCCTATGACCACTTCCAACCTCTTCCTCCCTGGGAA 454

Qy       1270 CATAATCCTTGGACTGCATGTTCCGTGTCCTGTGGAGGAGGGATTCAGAGACGGAGCTTT 1329
              ||||||||||||||||||||||||||| |||||||||||||||||||||||||||||| |
Db        455 CATAATCCTTGGACTGCATGTTCCGTGTCCTGTGGAGGAGGGATTCAGAGACGGAGCTTT 514

Qy       1330 GTGTGTGTAGAGGAATCCATGCATGGAGAGATACTGCAGGTGGAAGAATGGAAGTGCATC 1389
              ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        515 GTGTGTGTAGAGGAATCCATGCATGGAGAGATACTGCAGGTGGAAGAATGGAAGTGCATC 574

Qy       1390 TACCCACCCAAACCCAAGCCCATGCAAACTTGTAATCTGTTTGATTGCCCCAATGGGATT 1449
              ||||||||||||||||||   ||||||| ||||| |||| ||||||| |  |   |   |
Db        575 TACCCACCCAAACCCAAGGTTATGCAAACTTGTAATCTGTTTGATTGCCCCAATGGGATT 634

Qy       1450 GCCATGGAGTGGTCTCAGTGCACAGTGACTTGTGGCCGAGGGTTACGGTACCGGGTTGTT 1509
              |||||||||||||||||||||||||||| |||||||||||||||| |||| ||| |||
Db        635 GCCATGGAGTGGTCTCAGTGCACAGTGACTTGTGGCCGAGGGTTACGGTACCGGGTTGTT 694

Qy       1510 CTGTGTATTAACCACCCGGGAGAGCATGTTGGGGGCTGCAATCCACAACTGAAGTTACAC 1569
              ||||||||||||||  |  | |||||||||||||||||||||||||||||||| | | |
Db        695 CTGTGTATTAACCACCGCGGAGAGCATGTTGGGGGCTGCAATCCACAACTGAAGTTACAC 754

Qy       1570 ATCAAAGAAGAATGTGTCATCCCCATCCCGGGTTATAAACCAAAAGAAAAAAGTCCAGCC 1629
              ||||||||||||||||||   ||||||||||||||||||||||||||||||| |||  |
Db        755 ATCAAAGAAGAATGTGTCATCCCCATCCCGGGTTATAAACCAAAAGAAAAAAGTCCAGCC 814

Qy       1630 GAAGCAAAATTGCCTTGGCTGAAACAAGCACAAGAACTAGAAGAGACCAGAATAGCAACA 1689
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  |
Db        815 GAAGCAAAATTGCCTTGGCTGAAACAAGCACAAGAACTAGAAGAGACCAGAATAGCAACA 874

Qy       1690 GAAGAACCAACGTTCATTCCAGAACCCTGGTCAGCCTGCAGTACCACGTGTGGGCCAGGT 1749
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||    |
Db        875 GAAGAACCAACGTTCATTCCAGAACCCTGGTCAGCCTGCAGTACCACGTGTGGGCCGGGT 934

Qy       1750 GTGCAGGTCCGTGAGGTGAAGTGCCGTGTGGCTCCTCACATTCACGCAGACTGAGACTGAG 1809
              |||||||||  ||||||| |||||||||||||||||||||||||||||||||||||  |
Db        935 GTGCAGGTCCGTGAGGTGAAGTGCCGTGTGGCTCCTCACATTCACGCAGACTGAGACTGAG 994

Qy       1810 CTGCCCGAGGAAGAGTGTGAAGGCCCCAAGCTGCCCACCGAACGGCCCTGCCTCCTGGAA 1869
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        995 CTGCCCGAGGAAGAGTGTGAAGGCCCCAAGCTGCCCACCGAACGGCCCTGCCTCCTGGAA 1054

Qy       1870 GCATTTGATGAGAGCCCCGTTCCCCAGAGCTAGACATCCCTCTCCTGAGGACATTGACT 1929
              |||||||||||||||||    ||||||||||||||||||||||||||||||||| ||||
Db       1055 GCATTTGATGAGAGCCCCGTCCCCCAGAGCTAGACATCCCTCTCCTGAGGACAGTGAGT 1114

Qy       1930 ACGATTTACGACTCGGAGTACCCTGGGTTCACCCCTTGCACAGCAACATGCTGGGACGC 1989
               ||||| ||||| ||  ||| |||||||||||||||||||||| |||| |   |||||
Db       1115 ACGATTTACGACTCGGAGGACCCTGGGTTCACCCCTTGCACAGCAACATCCTGGGACGC 1174

Qy       1990 CATCAAGAAGCCATAGCAGTGGCTTACATATCCAGACCCAGCAGACAGTCAATGACAGT 2049
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||| ||
Db       1175 CATCAAGAAGCCATAGCAGTGGCTTACATATCCAGACCCAGCAGACAGTCAATGACAGT 1234

Qy       2050 TTGTTGGATATGGTCCACCTCCTCCAGCCATGAGCCAGGCCTGTAACACAGAGCCCTGT 2109
              |||||||||||||||| |   | |||||||||||||||||||||||||     |||  |
Db       1235 TTGTTGGATATGGTCCACCGTCCTCCAGCCATGAGCCAGGCCTGTAACACAGAGCCCTGT 1294

Qy       2110 CCCCCAGGTGGCATGTGGGCTCTTGGGGGCCCTGCTCAGCTACCTGTGGAGTTGGAATT 2169
              ||||||||||||||| |||||||||||||||||||||||||||||||||||||||||| |
Db       1295 CCCCCAGGTGGCATGTGGGCTCTTGGGGGCCCTGCTCAGCTACCTGTGGAGTTGGAATT 1354

Qy       2170 CAGACCCGAGATGTGTACTCCCTGCACCCAGGGGAGACCCCTGCCCCTCCTGAGGAGTGC 2229
              |||||||||||||||||  |  ||||||||||||||||||||||||| | ||| | | |
Db       1355 CAGACCCGAGATGTGTACTCCCTGCACCCAGGGGAGACCCCTGCCCCTCCTGAGGAGTGC 1414

Qy       2230 CGAGATGAAAAGCCCCATTCCTTACAAGCATGCAATCAGTTTGACTGCCCTCCTGGCTGG 2289
              ||||||||||||||| |     ||||||||||| |||||||||  ||||  |  |  |
Db       1415 CGAGATGAAAAGCCCCATCCCTTACAAGCATGCAATCAGTTTGACTGCCCTCCTGGCTGG 1474

Qy       2290 CACACTGAAGAATGGCAGCAGTGTTCCAGGACTTGTGGGGGGGGAACTCAGAACAGAACA 2349
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1475 CACATTGAAGAATGGCAGCAGTGTTCCAGGACTTGTGGGGGGGGAACTCAGAACAGAAGA 1534

Qy       2350 GTCACCTGTCGGCAGCTGCTAACGGATGGCAGCTTTTGAATCTCTCAGATGAATTGTGC 2409
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1535 GTCACCTGTCGGCAGCTGCTAACGGATGGCAGCTTTTGAATCTCTCAGATGAATTGTGC 1594

Qy       2410 CAAGGACCCAAGGCATCGTCTCACAAGTCCTGTGCCAGGACAGACTGTCCTCCACATTTA 2469
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db       1595 CAAGGACCCAAGGCATCGTCTCACAAGTCCTGTGCCAGGACAGACTGTCCTCCACATTTA 1654

Qy       2470 GCTGTGGGAGACTGGTCGAAGTGTTCTGTCAGTTGTGGTGTTGGAATCCAGAGAAGAAAG 2529
              ||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||
Db       1655 GCTGTGGGAGACTGGTCGAAGTGTTCTGTCAGTTGTGGTGTTGGAATCCAGAGAAGAAAG 1714

Qy       2530 CAGGTGTGTCAAAGGCTGGCAGCCAAAGGTCGGGCATCCCCCTCAGTGAGATGATGTGC 2589
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1715 CAGGTGTGTCAAAGGCTGGCAGCCAAAGGTCGGGCATCCCCCTCAGTGAGATGATGTGC 1774

Qy       2590 AGGGATCTACCAGGCTTCCTCTTGTAAGATCTTGCCAGATGCCTGAGTGCAGTAAAATC 2649
              ||||||||||||| |||| ||||||||||||||||||||| |  |  |  | ||||||||
Db       1775 AGGGATCTACCAGGGTTCCTTTTGTAAGATCTTGCCAGATGCCTGAGTGCAGTAAAATC 1834

Qy       2650 AAATCAGAGATGAAGACAAAACTTGGTGAGCAGGGTCCGCAGATCCTCAGTGTCCAGAGA 2709
              |||||||||||||||||||||||||||||||||||||||||| ||||| |||||||||||
Db       1835 AAATCAGAGATGAAGACAAAACTTGGTGAGCAGGGTCCGCAGATCCTCAGTGTCCAGAGA 1894

Qy       2710 GTCTACATTCAGACAAGGGAAGAGAAGCGTATTAACCTGACCATTGGTAGCAGAGCCTAT 2769
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1895 GTCTACATTCAGACAAGGGAAGAGAAGCGTATTAACCTGACCATTGGTAGCAGAGCCTAT 1954

Qy       2770 TTGCTGCCCAACACATCCGTGATTATTAAGTGCCCAGTGCGACGATTCCAGAAATCTCTG 2829
              |||||||||||||||||||||||||||||||||||| ||   ||||  ||||||||||||
Db       1955 TTGCTGCCCAACACATCCGTGATTATTAAGTGCCCAGTGCGACGATTCCAGAAATCTCTG 2014

Qy       2830 ATCCAGTGGGAGAAGGATGGCCGTTGCCTGCAGAACTCCAAACGGCTTGGCATCACCAAG 2889
              |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db       2015 ATCCAGTGGGAGAAGGATGGCCGTTGCCTGCAGAACTCCAAACGGCTTGGCATCACCAAG 2074

Qy       2890 TCAGGCTCACTAAAAATCCACGGTCTTGCTGCCCCGACATCGGCGTGTACCGGTGCATT 2949
              |||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db       2075 TCAGGCTCACTAAAAATCCACGGTCTTGCTGCCCCGACATCGGCGTGTACCGGTGCATT 2134

Qy       2950 GCAGGCTCTGCACAGGAAACAGTTGTGCTCAAGCTCATTGGTACTGACAACCGGCTCATC 3009
              |||||||||||||||||||| |||||||||||||||||||    | ||| |  |||||||
Db       2135 GCAGGCTCTGCACAGGAAACAGTTGTGCTCAAGCTCATTGGTACTGACAACCGGCTCATC 2194

Qy       3010 GCACGCCCAGCCCTCAGGGAGCTATGAGGGAATATCCTGGGATGGACCACAGCGAAGCC 3069
              ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2195 GCACGCCCAGCCCTCAGGGAGCTATGAGGGAATATCCTGGGATGGACCACAGCGAAGCC 2254

Qy       3070 AATAGTTTGGGAGTCACATCCCACAAAAATGACCCAAATGTGAATAACAAAAATGACCTT 3129
              |||||||||||||||||||||| |||| || ||| |||| ||| ||||  | ||||||
Db       2255 AATAGTTTGGGAGTCACATCCCACAAAAATGACCCAAATGTGAATAACAAAAATGACCTT 2314

Qy       3130 TATCTGGATGATGACCACATTAGTAACCAGCCTTTCTTGAGAGCTCTTGTTAGGCTACTGC 3189
              |||||||||||||||||||||||||||||||||||||||       ||||| |||||
Db       2315 TATCTGGATGATGACCACATTAGTAACCAGCCTTTCTTGAGAGCTCTTGTTAGGCTACTGC 2374

Qy       3190 AGCAATTCTGCAGGAAGCACCAACTCCTGGGAGTTGAAGAATAAGCAGTTGGAAGCAGCA 3249
              ||||||||||||||||||||||||||||||||||||||||||| ||||| | ||||||||
Db       2375 AGCAATTCTGCAGGAAGCACCAACTCCTGGGAGTTGAAGAATAAGCAGTTGGAAGCAGCA 2434

Qy       3250 GTTAAACAAGGACCATATAGCATGGATACAGCAATGCATGACCTGATAAGAAACATG 3309
              |||||||||| ||||||||||||||||||||||||| |||||||  ||||||||||
Db       2435 GTTAAACAAGGACCATATAGCATGGATACAGCAATGCATGACCTGATAAGAAACATG 2494

Qy       3310 AGTCAGCTCATGGAAACCGGAGAGGTCAGCGATGATCTTCCTCCTCAGCTGATATATCAG 3369
              ||||||||||||||||||| |||||||||||||||||||||||||||||||| |||||||
Db       2495 AGTCAGCTCATGGAAACCGGAGAGGTCAGCGATGATCTTCCTCCTCAGCTGATATATCAG 2554

Qy       3370 CTGGTGGCCGAATTAGCCAAGGCACAGCCAACACACATGCATCCCCCGGGCATCCAGGAA 3429
              ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db       2555 CTGGTGGCCGAATTAGCCAAGGCACAGCCAACACACATGCATCCCCCGGGCATCCAGGAA 2614
```

```
Qy       3430 GAGACACCTCCTGCTGCTCAGCTCAGAGGGGAAACAGGGAGTGTGTCCCAAAGCTCGCAT 3489
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2615 GAGACACCTCCTGCTGCTCAGCTCAGAGGGGAAACAGGGAGTGTGTCCCAAAGCTCGCAT 2674

Qy       3490 GCAAAAAACTCAGGCAAGCTGACATTCAAGCCGAAAGGACCTGTTCTCATGAGGCAAAGC 3549
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2675 GCAAAAAACTCAGGCAAGCTGACATTCAAGCCGAAAGGACCTGTTCTCATGAGGCAAAGC 2734

Qy       3550 CAACCTCCCTCAATTTCATTTAATAAAACAATAAATTCCAGGATTGGAAATACAGTATAC 3609
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2735 CAACCTCCCTCAATTTCATTTAATAAAACAATAAATTCCAGGATTGGAAATACAGTATAC 2794

Qy       3610 ATTACAAAAAGGACAGAGGTCATCAATATACTGTGTGACCTTATTACCCCCAGTGAGGCC 3669
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2795 ATTACAAAAAGGACAGAGGTCATCAATATACTGTGTGACCTTATTACCCCCAGTGAGGCC 2854

Qy       3670 ACATATACATGGACCAAGGATGGAACCTTGTTACAGCCCTCAGTAAAAATAATTTTGGAT 3729
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2855 ACATATACATGGACCAAGGATGGAACCTTGTTACAGCCCTCAGTAAAAATAATTTTGGAT 2914

Qy       3730 GGAACTGGGAAGATACAGATACAGAATCCTACAAGGAAAGAACAAGGCATATATGAATGT 3789
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2915 GGAACTGGGAAGATACAGATACAGAATCCTACAAGGAAAGAACAAGGCATATATGAATGT 2974

Qy       3790 TCTGTAGCTAATCATCTTGGTTCAGATGTGGAAAGTTCTTCTGTGCTGTATGCAGAGGCA 3849
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2975 TCTGTAGCTAATCATCTTGGTTCAGATGTGGAAAGTTCTTCTGTGCTGTATGCAGAGGCA 3034

Qy       3850 CCTGTCATCTTGTCTGTTGAAAGAAATATCACCAAACCAGAGCACAACCATCTGTCTGTT 3909
              |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db       3035 CCTGTCATCTTGTCTGTTGAAAGAAATATCACCAAACCAGAGCACAACCATCTGTCTGTT 3094

Qy       3910 GTGGTTGGAGGCATCGTGGAGGCAGCCCTTGGAGCAAACGTGACAATCCGATGTCCTGTA 3969
              |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db       3095 GTGGTTGGAGGCATCGTGGAGGCAGCCCTTGGAGCAAACGTGACAATCCGATGTCCTGTA 3154

Qy       3970 AAAGGTGTCCCTCAGCCTAATATAACTTGGTTGAAGAGAGGAGGATCTCTGAGTGGCAAT 4029
              |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db       3155 AAAGGTGTCCCTCAGCCTAATATAACTTGGTTGAAGAGAGGAGGATCTCTGAGTGGCAAT 3214

Qy       4030 GTTTCCTTGCTTTTCAATGGATCCCTGTTGTTGCAGAATGTTTCCCTTGAAAATGAAGGA 4089
              |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db       3215 GTTTCCTTGCTTTTCAATGGATCCCTGTTGTTGCAGAATGTTTCCCTTGAAAATGAAGGA 3274

Qy       4090 ACCTACGTCTGCATAGCCACCAATGCTCTTGGAAAGGCAGTGGCAACATCTGTACTCCAC 4149
              |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db       3275 ACCTACGTCTGCATAGCCACCAATGCTCTTGGAAAGGCAGTGGCAACATCTGTACTCCAC 3334

Qy       4150 TTGCTGGAACGAAGATGGCCAGAGAGTAGAATCGTATTTCTGCAAGGACATAAAAAGTAC 4209
              |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db       3335 TTGCTGGAACGAAGATGGCCAGAGAGTAGAATCGTATTTCTGCAAGGACATAAAAAGTAC 3394

Qy       4210 ATTCTCCAGGCAACCAACACTAGAACCAACAGCAATGACCCAACAGGAGAACCCCCGCCT 4269
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3395 ATTCTCCAGGCAACCAACACTAGAACCAACAGCAATGACCCAACAGGAGAACCCCCGCCT 3454

Qy       4270 CAAGAGCCTTTTTGGGAGCCTGGTAACTGGTCACATTGTTCTGCCACCTGTGGTCATTTG 4329
              |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db       3455 CAAGAGCCTTTTTGGGAGCCTGGTAACTGGTCACATTGTTCTGCCACCTGTGGTCATTTG 3514

Qy       4330 GGAGCCCGCATTCAGAGACCCCAGTGTGTGATGGCCAATGGGCAGGAAGTGAGTGAGGCC 4389
              |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db       3515 GGAGCCCGCATTCAGAGACCCCAGTGTGTGATGGCCAATGGGCAGGAAGTGAGTGAGGCC 3574

Qy       4390 CTGTGTGATCACCTCCAGAAGCCACTGGCTGGGTTTGAGCCCTGTAACATCCGGGACTGC 4449
              |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db       3575 CTGTGTGATCACCTCCAGAAGCCACTGGCTGGGTTTGAGCCCTGTAACATCCGGGACTGC 3634

Qy       4450 CCAGCGAGGTGGTTCACAAGTGTGTGGTCACAGTGCTCTGTGTCTTGCGGTGAAGGATAC 4509
              |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db       3635 CCAGCGAGGTGGTTCACAAGTGTGTGGTCACAGTGCTCTGTGTCTTGCGGTGAAGGATAC 3694

Qy       4510 CACAGTCGGCAGGTGACGTGCAAGCGGACAAAAGCCAATGGAACTGTGCAGGTGGTGTCT 4569
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3695 CACAGTCGGCAGGTGACGTGCAAGCGGACAAAAGCCAATGGAACTGTGCAGGTGGTGTCT 3754

Qy       4570 CCAAGAGCATGTGCCCCTAAAGACCGGCCTCTGGGAAGAAAACCATGTTTTGGTCATCCA 4629
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3755 CCAAGAGCATGTGCCCCTAAAGACCGGCCTCTGGGAAGAAAACCATGTTTTGGTCATCCA 3814

Qy       4630 TGTGTTCAGTGGGAACCAGGGAACCGGTGTCCTGGACGTTGCATGGGCCGTGCTGTGAGG 4689
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3815 TGTGTTCAGTGGGAACCAGGGAACCGGTGTCCTGGACGTTGCATGGGCCGTGCTGTGAGG 3874

Qy       4690 ATGCAGCAGCGTCACACAGCTTGTCAACACAACAGCTCTGACTCCAACTGTGATGACAGA 4749
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3875 ATGCAGCAGCGTCACACAGCTTGTCAACACAACAGCTCTGACTCCAACTGTGATGACAGA 3934

Qy       4750 AAGAGACCCACCTTAAGAAGGAACTGCACATCAGGGGCCTGTGATGTGTGTTGGCACACA 4809
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3935 AAGAGACCCACCTTAAGAAGGAACTGCACATCAGGGGCCTGTGATGTGTGTTGGCACACA 3994

Qy       4810 GGCCCTTGGAAGCCCTGTACAGCAGCCTGTGGCAGGGGTTTCCAGTCTCGGAAAGTCGAC 4869
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3995 GGCCCTTGGAAGCCCTGTACAGCAGCCTGTGGCAGGGGTTTCCAGTCTCGGAAAGTCGAC 4054

Qy       4870 TGTATCCACACAAGGAGTTGCAAACCTGTGGCCAAGAGACACTGTGTACAGAAAAAGAAA 4929
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4055 TGTATCCACACAAGGAGTTGCAAACCTGTGGCCAAGAGACACTGTGTACAGAAAAAGAAA 4114

Qy       4930 CCAATTTCCTGGCGGCACTGTCTTGGGCCCTCCTGTGATAGAGACTGCACAGACACAACT 4989
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4115 CCAATTTCCTGGCGGCACTGTCTTGGGCCCTCCTGTGATAGAGACTGCACAGACACAACT 4174

Qy       4990 CACTACTGTATGTTTGTAAAACATCTTAATTTGTGTTCTCTAGACCGCTACAAACAAAGG 5049
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4175 CACTACTGTATGTTTGTAAAACATCTTAATTTGTGTTCTCTAGACCGCTACAAACAAAGG 4234

Qy       5050 TGCTGCCAGTCATGTCAAGAGGGATAA 5076
              |||||||||||||||||||||||||||
Db       4235 TGCTGCCAGTCATGTCAAGAGGGATAA 4261


RESULT 6
AAD14372
ID   AAD14372 standard; DNA; 8578 BP.
XX
AC   AAD14372;
XX
DT   01-NOV-2001  (first entry)
XX
DE   Novel human protein gene open reading frame with flanking sequence.
XX
KW   Novel human protein; NHP; thrombospondin; gene therapy; cancer;
KW   cytostatic; antisense therapy; angiogenesis; biological disorder; ds.
XX
OS   Homo sapiens.
XX
PN   WO200161011-A2.
XX
PD   23-AUG-2001.
XX
PF   15-FEB-2001; 2001WO-US05290.
XX
PR   17-FEB-2000; 2000US-0183282.
XX
PA   (LEXI-) LEXICON GENETICS INC.
XX
PI   Donoho G,  Scoville J,  Turner CA,  Friedrich G,  Zambrowicz B;
PI   Sands AT;
XX
DR   WPI; 2001-514776/56.
XX
PT   An isolated nucleic acid encoding a thrombospondin useful as a
PT   hybridization probe and gene therapy treatments of cancer -
```

```
XX
PS   Disclosure; Page 53-56; 56pp; English.
XX
CC   The invention relates to novel human polynucleotides encoding proteins
CC   that share sequence similarity with animal proteins having thrombospondin
CC   repeats. NHPs are expressed in human cell lines, pituitary, lymph nodes,
CC   prostrate, testis, adrenal gland, uterus, foetal kidney, foetal lung and
CC   gene trapped human cells. NHPs are thrombospondins useful for treating
CC   biological disorders involving angiogenesis, cancer and development and
CC   also in pharmacogenomic applications. NHPs are useful as a hybridisation
CC   probe for screening libraries, assessing gene expression patterns and
CC   also in gene therapy. Proteins having thrombospondin repeats act as
CC   receptors, secreted extracellular matrix proteins and proteases. The
CC   present sequence is novel human protein (NHP) gene containing NHP
CC   open reading frame with flanking sequences.
XX
SQ   Sequence 8578 BP; 2449 A; 1827 C; 2076 G; 2241 T; 0 other;

  Query Match             71.0%;  Score 3755;  DB 22;  Length 8578;
  Best Local Similarity   79.7%;  Pred. No. 0;
  Matches 5076;  Conservative    0;  Mismatches    0;  Indels 1291;  Gaps    [illegible];

Qy          1 ATGGCTTGCTGGAGGAG[illegible]TGATAGGGATG[illegible]ATG A[illegible] 60
                ||| |||||||||| |||    |||| |||||||||||||| |||  | |
Db        179 ATGGCTTGCTGGAGGAG[illegible]GGGGCTGATAGGGATGG[illegible]ATG A[illegible] 238

Qy         61 [illegible]CAGACCACA[illegible]GAGAAA[illegible]GGAG[illegible]TATTT[illegible]GGAGTT[illegible]CA[illegible] 120
              |||||| ||||||||      ||||||||||||||||||| ||||||||||  |
Db        239 CTCCCGCAGACCACAGTGAGAAATCTCTGGAGCCTATTTCCTCCCGAGTTGGCATTT 298

Qy        121 TCTCCTCAGGGAAGTTTTCTGGAAGACACAACAGGGGAGCAGTTCCTCACTTATCGCTAT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        299 TCTCCTCAGGGAAGTTTTCTGGAAGACACAACAGGGGAGCAGTTCCTCACTTATCGCTAT 358

Qy        181 GATGACCAGACCTCAAGAAACACTCGTTCAGATGAAGACAAAGATGGCAATCTGGATCCT 240
              ||||||||||||||||||||| ||||| |||||||||||||||||||||||||| |  |
Db        359 GATGACCAGACCTCAAGAAACACTCGTTCAGATGAAGACAAAGATGGCAATCTGGATCCT 418

Qy        241 TGGGGCGACTGGAGTGACTGCTCCCGGACCTGTGGGGGAGGAGCATCATATTCTCTGCGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        419 TGGGGCGACTGGAGTGACTGCTCCCGGACCTGTGGGGGAGGAGCATCATATTCTCTGCGG 478

Qy        301 AGATGTTTGACTGGAAGGAATTGTGAAGGGCAGAACATTCGGTACAAGACATGCAGCAAT 360
              |||||||||||||||||| ||||||||||||||||||||||||||||||||| |||| ||
Db        479 AGATGTTTGACTGGAAGGAATTGTGAAGGGCAGAACATTCGGTACAAGACATGCAGCAAT 538

Qy        361 CATGACTGCCCTCCAGATGCAGAAGATTTCAGAGCCCAGCAGTGCTCAGCCTACAATGAT 420
              ||||| ||||||||||||| |||||||||||||||||||||||||||||||||||| ||
Db        539 CATGACTGCCCTCCAGATGCAGAAGATTTCAGAGCCCAGCAGTGCTCAGCCTACAATGAT 598

Qy        421 GTCCAGTATCAGGGGCATTACTATGAATGGCTTCCACGATATAATGATCCTGCTGCCCCG 480
              ||||||||||||||||||||||||||||||||||||||||||||| |  ||||||| ||
Db        599 GTCCAGTATCAGGGGCATTACTATGAATGGCTTCCACGATATAATGATCCTGCTGCCCCG 658

Qy        481 TGTGCACTCAAGTGTCATGCAGAAGGACAAAACTTGGTGGTGGAGCTGGCACCTAAGGTA 540
              |||| |||||||||||||||   |||||||||||||||||||||||||||| ||| |
Db        659 TGTGCACTCAAGTGTCATGCAGAAGGACAAAACTTGGTGGTGGAGCTGGCACCTAAGGTA 718

Qy        541 CTGGATGGAACTCGTTGCAACACGGACTCCTTGGACATGTGTATCAGTGGCATCTGTCAG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        719 CTGGATGGAACTCGTTGCAACACGGACTCCTTGGACATGTGTATCAGTGGCATCTGTCAG 778

Qy        601 GCAGTGGGCTGCGATCGGCAACTGGGAAGCAATGCCAAGGAGGACAACTGTGGAGTCTGT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        779 GCAGTGGGCTGCGATCGGCAACTGGGAAGCAATGCCAAGGAGGACAACTGTGGAGTCTGT 838

Qy        661 GCCGGCGATGGCTCCACCTGCAGGCTTGTACGGGGACAATCAAAGTCACACGTTTCTCCT 720
              ||| ||||||||| |||||   ||||||||||||||||||||||||||||||||||||||
Db        839 GCCGGCGATGGCTCCACCTGCAGGCTTGTACGGGGACAATCAAAGTCACACGTTTCTCCT 898

Qy        721 GAAAAAAGAGAAGAAAATGTAATTGCTGTTCCTTGGGAAGTCGAAGTGTGAGAATTACA 780
              ||| ||||||| | |||||||||||||||||||||||||||||||| | ||||||||
Db        899 GAAAAAAGAGAAGAAAATGTAATTGCTGTTCCTTGGGAAGTCGAAGTGTGAGAATTACA 958

Qy        781 GTGAAAGGACCTGCCCACCTCTTTATTGAATCAAAAACACTTCAAGGAAGCAAAGGAGAA 840
              |||||||||||||||||||||||||||| | ||||||| |||| | | | ||||||||
Db        959 GTGAAAGGACCTGCCCACCTCTTTATTGAATCAAAAACACTTCAAGGAAGCAAAGGAGAA 1018

Qy        841 CACAGCTTTAACAGCCCTGGCGTCTTTGTCGTAGAAAACACAACAGTGGAATTTCAGAGG 900
              |||||||||||||||||||||||||||||||||||||| ||| |||||||||||||||
Db       1019 CACAGCTTTAACAGCCCTGGCGTCTTTGTCGTAGAAAACACAACAGTGGAATTTCAGAGG 1078

Qy        901 GGCTCCGAGAGGCAAACTTTTAAGATTCCAGGACCTCTGATGGCTGATTTCATCTTCAAG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1079 GGCTCCGAGAGGCAAACTTTTAAGATTCCAGGACCTCTGATGGCTGATTTCATCTTCAAG 1138

Qy        961 ACCAGGTACACTGCAGCCAAAGACAGCGTGGTTCAGTTCTTCTTTTACCAGCCCATCAGT 1020
              ||||||||||||||||||||||||||||||    |   |  ||||||||||||||||||
Db       1139 ACCAGGTACACTGCAGCCAAAGACAGCGTGGTTCAGTTCTTCTTTTACCAGCCCATCAGT 1198

Qy       1021 CATCAGTGGAGACAAACTGACTTCTTTCCCTGCACTGTGACGTGTGGAGGAGGTTATCAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1199 CATCAGTGGAGACAAACTGACTTCTTTCCCTGCACTGTGACGTGTGGAGGAGGTTATCAG 1258

Qy       1081 CTCAATTCTGCTGAATGTGTGGATATCCGCTTGAAGAGGGTAGTTCCTGACCATTATTGT 1140
              ||||||||||||||||||||||||| |||||||||||||||| ||||||||||||||||
Db       1259 CTCAATTCTGCTGAATGTGTGGATATCCGCTTGAAGAGGGTAGTTCCTGACCATTATTGT 1318

Qy       1141 CACTACTACCCTGAAAATGTAAAACCAAAACCAAAACTGAAGGAATGCAGCATGGATCCC 1200
              |||||||||| ||||||||||||||||      |     |  ||||||||||||||||||
Db       1319 CACTACTACCCTGAAAATGTAAAACCAAAACCAAAACTGAAGGAATGCAGCATGGATCCC 1378

Qy       1201 TGCCCATCAAGTGATGGATTTAAAGAGATAATGCCCTATGACCACTTCCAACCTCTTCCT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1379 TGCCCATCAAGTGATGGATTTAAAGAGATAATGCCCTATGACCACTTCCAACCTCTTCCT 1438

Qy       1261 CGAGCTGGGAACATAATCCTTGGACTGCATGGTCGTTCCTCGTCGGAGGAGGATTCAGA 1318
              |  |||||||||||||||||||||||||||||||| |||| |   ||||||||||||||
Db       1439 CGAGCTGGGAACATAATCCTTGGACTGCATGGTCGTTCCTCGTCGGAGGAGGATTCAGA 1498

Qy       1319 GACGGAGCTTTGTGTGTGTAGAGGAATCCATGCATGGAGAGATATTGCAGGTGGAAGAAT 1378
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1499 GACGGAGCTTTGTGTGTGTAGAGGAATCCATGCATGGAGAGATATTGCAGGTGGAAGAAT 1558

Qy       1379 GGAAGTGCATGTACGCACCCAAACCCAAGGTTATGCAAACTTGTAATCTGTTTGATTGCC 1438
              |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db       1559 GGAAGTGCATGTACGCACCCAAACCCAAGGTTATGCAAACTTGTAATCTGTTTGATTGCC 1618

Qy       1439 CCAAGTGGATTGCCATGGAGTGGTCTCAGTGCACAGTGACTTGTGGGAGGGTTACGGT 1498
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1619 CCAAGTGGATTGCCATGGAGTGGTCTCAGTGCACAGTGACTTGTGGGAGGGTTACGGT 1678

Qy       1499 ACCGGGTTGTTCTGTGTATTAACCACCGGGAGAGCATGTGGGGCTGCAATCCACAAC 1558
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1679 ACCGGGTTGTTCTGTGTATTAACCACCGGGAGAGCATGTGGGGCTGCAATCCACAAC 1738

Qy       1559 TGAAGTTACACATCAAAGAACAATGTGTCATTCCCATCCCGCTTATAAACCAAAAGAAA 1618
              ||||||||||||||||||||||||||||||||||||||| | ||||||||||||||||
Db       1739 TGAAGTTACACATCAAAGAACAATGTGTCATTCCCATCCCGCTTATAAACCAAAAGAAA 1798

Qy       1619 AAAGTCCAGTGGAAGCAAAATTGCCTTGGCTGAAACAAGCACAAGAACTAGAAGAGACCA 1678
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1799 AAAGTCCAGTGGAAGCAAAATTGCCTTGGCTGAAACAAGCACAAGAACTAGAAGAGACCA 1858

Qy       1679 GAATAGCAACAGAAGAACCAACGTTCATTCCAGAACCCTGGTTCAGCCTGCAGTACCAGGT 1738
              ||||||| ||||||||||||||||||||||||||||||| |||||||||||||||||||
Db       1859 GAATAGCAACAGAAGAACCAACGTTCATTCCAGAACCCTGGTTCAGCCTGCAGTACCAGGT 1918

Qy       1739 GTGGGCCAGGTGTGCAGGTTCGGCGAGGTGAAGTGCCTGTGCTCCTTCACATTCACGCAGA 1798
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1919 GTGGGCCAGGTGTGCAGGTTCGGCGAGGTGAAGTGCCTGTGCTCCTTCACATTCACGCAGA 1978

Qy       1799 CTGAGACTGAGCTGCCTGAGGAAGAGTGTGAAGGCCCCAAGCTGCCCACCGAACGGCCCT 1858
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1979 CTGAGACTGAGCTGCCTGAGGAAGAGTGTGAAGGCCCCAAGCTGCCCACCGAACGGCCCT 2038
```

```
Qy   1859 GCCTCCTGGAAGCATGTGATGAGAGCCCGGCCTCCCGAGAGCTAGACATCCCTCTCCCTG 1918
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2039 GCCTCCTGGAAGCATGTGATGAGAGCCCGGCCTCCCGAGAGCTAGACATCCCTCTCCCTG 2098

Qy   1919 AGGACAGTGAGACGACTTACGACTGGGAGTACGCTGGGTTCACCCCTTGCACAGCAACAT 1978
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2099 AGGACAGTGAGACGACTTACGACTGGGAGTACGCTGGGTTCACCCCTTGCACAGCAACAT 2158

Qy   1979 GCTTGGGAGGCCATCAAGAAGCCATAGCAGTGTGCTTACATATCCAGACCCAGCAGACAG 2038
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2159 GCTTGGGAGGCCATCAAGAAGCCATAGCAGTGTGCTTACATATCCAGACCCAGCAGACAG 2218

Qy   2039 TCAATGACAGCTTGTGTGATATGGTCCACCGTCCTCCAGCCATGAGCCAGGCCTGTAACA 2098
          |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db   2219 TCAATGACAGCTTGTGTGATATGGTCCACCGTCCTCCAGCCATGAGCCAGGCCTGTAACA 2278

Qy   2099 CAGAGCCCTGTCCCCCCAGG---------------------------------------- 2118
          ||||||||||||||||||||
Db   2279 CAGAGCCCTGTCCCCCCAGGAGAGAGCCAGCAGCTTGTAGAAGCATGCCGGGTTACATAA 2338

Qy   2119 ------------------------------------------------------------ 2118

Db   2339 TGGTCCTGCTAGTCTGAGGAGAGCCTTCTTCTCTAACAGGATTCAACACTGCTAGGGAAG 2398

Qy   2119 ------------------------------------------------------------ 2118

Db   2399 AAAGGAGGAAAGCAAGAGGCAATAGTGATGTGTTTCTGTACCAGCTTGTTACCTATTTCT 2458

Qy   2119 ------------------------------------------------------------ 2118

Db   2459 TGATATAAAAAACAATTCTTTATTGAGTTCATTGTCTGTGAATAAGAAATTGTTGCCCAT 2518

Qy   2119 ---------------------------------------------TGGCATGTGGGCTCT 2133
                                                       |||||||||||||||
Db   2519 TTCTTAAATAAAAACAGCTCCATCTCCAAAAAAAAAAAAAAAAAATGGCATGTGGGCTCT 2578

Qy   2134 TGGGGGCCCTGCTCAGCTACCTGTGGAGTTGGAATTCAGACCCGAGATGTGTACTGCCTG 2193
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2579 TGGGGGCCCTGCTCAGCTACCTGTGGAGTTGGAATTCAGACCCGAGATGTGTACTGCCTG 2638

Qy   2194 CACCCAGGGGAGACCCCTGCCCCTCCTGAGGAGTGCCGAGATGAAAAGCCCCATGCTTTA 2253
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2639 CACCCAGGGGAGACCCCTGCCCCTCCTGAGGAGTGCCGAGATGAAAAGCCCCATGCTTTA 2698

Qy   2254 CAAGCATGCAATCAGTTTGACTGCCCTCCTGGCTGGCACATTGAAGAATGGCAGCAGTGT 2313
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2699 CAAGCATGCAATCAGTTTGACTGCCCTCCTGGCTGGCACATTGAAGAATGGCAGCAGTGT 2758

Qy   2314 TCCAGGACTTGTGGCGGGGGAACTCAGAACAGAAGAGTCACCTGTCGGCAGCTGCTAACG 2373
          ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db   2759 TCCAGGACTTGTGGCGGGGGAACTCAGAACAGAAGAGTCACCTGTCGGCAGCTGCTAACG 2818

Qy   2374 GATGGCAGCTTTTTGAATCTCTCAGATGAATTGTGCCAAGGACCCAAGGCATCGTCTCAC 2433
          ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db   2819 GATGGCAGCTTTTTGAATCTCTCAGATGAATTGTGCCAAGGACCCAAGGCATCGTCTCAC 2878

Qy   2434 AAGTCCTGTGCCAGGACAGACTGTCCTCCACATTTAGCTGTGGGAGACTGGTCGAAG--- 2490
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2879 AAGTCCTGTGCCAGGACAGACTGTCCTCCACATTTAGCTGTGGGAGACTGGTCGAAGGAG 2938

Qy   2491 ------------------------------------------------------------ 2490

Db   2939 CATTCAATGCAAGAGGACAATGGAGCAGGATCTACACAATTCTAAAGAAAAGCAAGCATG 2998

Qy   2491 ------------------------------------------------------------ 2490

Db   2999 ACTCAAGGATTTCCTCTTCATCCTGTGTTCTTCATGTAGAGAGACAGCAGAGAGGCAGTC 3058

Qy   2491 ------------------------------------------------------------ 2490

Db   3059 AGAGAATACTGTCTGATAAGCCCTTGAAAAAGCTGTAGGGCCAAGATGAGATACAGAGAT 3118

Qy   2491 ------------------------------------------------------------ 2490
```

```
Db   3119 GACTCAAAACAGAGAATCCAGGAATGCATAGATCCTGGTAAAAAGGTGGGAGATGAGTA 3178

Qy   2491 ------------------------------------------------------------ 2490

Db   3179 ATAAATTCATTTGTGTAGGATTAAGACTAATCAACTAACAATTATATTATAGAACATAAC 3238

Qy   2491 ------------------------------------------------------------ 2490

Db   3239 ATAAATATCAGAAATCTTGACATTATCTAAATAATAAAATGAAAACTAATTGAGATTTGG 3298

Qy   2491 ------------------------------------------------------------ 2490

Db   3299 AGAGATGAGGTAGATGATATAGTTTGGCTGTGTCCCCACCAAATCTCATCTTGAATTGT 3358

Qy   2491 ------------------------------------------------------------ 2490

Db   3359 AGTTCCCATAATTCCCATGTGTTGTGGAGGGACTCAGTTGGAGATAATTGAATCATGGG 3418

Qy   2491 ------------------------------------------------------------ 2490

Db   3419 GGCAGTTTCCCTCATACTGTTCTCGTGGTGGTAAATGAGTCTCACGAGATCTGATGGTTT 3478

Qy   2491 ------------------------------------------------------------ 2490

Db   3479 TATAAGGGGTTTCCCTTTTCGCTTGGCTCTCATTCTCTCTTGCCTGCTGCCATGTAAGAC 3538

Qy   2491 ------------------------------------------------------------ 2490

Db   3539 GTCCCTTTGCCCTTCCTTTGTCTTCTGCCATGATTGTGAGGCTTCCCTAGCCACGTGGAA 3598

Qy   2491 ------------------------------------------------------------ 2490

Db   3599 CTGAGTCTATTAAACCTCTTTCCTTTATAACTTACCCAGTCTTGGGTATGTCTTTATTAA 3658

Qy   2491 ------------------------------------------------------------ 2490

Db   3659 CAACATAAGATTGGACTAATACAGTAGAGGAAATGTAAGTGTGCTTATTTCCTCATCCTT 3718

Qy   2491 ------------------------------------------------------------ 2490

Db   3719 CTTAGTAGCAAGTCAATAAATACTCTCCTAAGTCAAATTGTCATTAAAAATAACTATCCA 3778

Qy   2491 ------------------------------------------------------------ 2490

Db   3779 AATCTCTTGTTGGTTTATTTAATCTTCTTTATTAACTTTAGAGTGTTCTTTCGGGAATTA 3838

Qy   2491 ------------------------------------------------------------ 2490

Db   3839 ATCATGGTTTAAAAAATATCAAACATTCAACAACTCTAATTTTACTTTAATGTCTTTTTT 3898

Qy   2491 ------------------------------------------------------------ 2490

Db   3899 TCTAATATATCTAATAAAATTGCATTAAATTTTTAAGTTGAAAAAAAAAAAAAAAAAAAA 3958

Qy   2491  TGTTCTGTCAGTTGTGGTGTTGGAATCCAGAGAAGAAAGCAGGTGTGTCAAAGGCTGGC 2549
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3959 ATGTTCTGTCAGTTGTGGTGTTGGAATCCAGAGAAGAAAGCAGGTGTGTCAAAGGCTGGC 4018

Qy   2550 AGCCAAAGGTCGGCGCATCCCCCTCAGTGAGATGATGTGCAGGGATCTACCAGGGTTCCC 2609
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4019 AGCCAAAGGTCGGCGCATCCCCCTCAGTGAGATGATGTGCAGGGATCTACCAGGGTTCCC 4078

Qy   2610 TCTTGTAAGATCTTGCCAGATGCCTGAGTGCAGTAAAATCAAATCAGAGATGAAGACAAA 2669
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4079 TCTTGTAAGATCTTGCCAGATGCCTGAGTGCAGTAAAATCAAATCAGAGATGAAGACAAA 4138

Qy   2670 ACTTGGTGAGCAGGGTCCGCAGATCCTCAGTGTCCAGAGAGTCTACATTCAGACAAGGGA 2729
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4139 ACTTGGTGAGCAGGGTCCGCAGATCCTCAGTGTCCAGAGAGTCTACATTCAGACAAGGGA 4198

Qy   2730 AGAGAAGCGTATTAACCTGACCATTGGTAGCAGAGCCTATTTGCTGCCCAACACATCCGT 2789
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4199 AGAGAAGGTATTAACTTGACCATTGGTAGCACAGCCTATTTGCTGCCCAACACATCCGT 4258

Qy       2790 GATTATTAAGTGCCCCTTGCGACGATTCCAGAAATCTCTGATCCAGTTGGAGAATGATCT 2849
              ||||||||||||||      |  ||||    ||||||||||||||||||      |  |
Db       4259 GATTATTAAGTGCCCCTTGCGACGATTCCAGAAATCTCTGATCCAGTTGGAGAAGGATCT 4318

Qy       2850 CCGTTGCCTGCAGAACTCCAAACGGCTTGGCATCACCAAGTCAGGCTCACTAAAAATCCA 2909
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4319 CCGTTGCCTGCAGAACTCCAAACGGCTTGGCATCACCAAGTCAGGCTCACTAAAAATCCA 4378

Qy       2910 CCGCTTGCTGGCCCCCGACACTCGGCGTGTACCGGCGCATTGCAGGCTCTGCACAGGAAAC 2969
              |||||||||||||    | ||||||| |||| ||||||||||           |
Db       4379 CCGCTTGCTGGCCCCCGACACTCGGCGTGTACCGGCGCATTGCAGGCTCTGCACAGGAAAC 4438

Qy       2970 ACTTCTTGCTCAAGCTCATTGTTACTGACAACCGGCTCATGGCACGGCCAGCCCTCAGGGA 3029
              ||||||||||||||||||||||||||||||||||||||||||| | |||||||||||||
Db       4439 ACTTCTTGCTCAAGCTCATTGTTACTGACAACCGGCTCATGGCACGGCCAGCCCTCAGGGA 4498

Qy       3030 GCCTATGAGGGAATATCCTCGGATGGACCACAGGAAGCAATACTTTGGAGTCACATC 3089
              ||||| ||||||||    |||||||||||||||||||||||| |
Db       4499 GCCTATGAGGGAATATCCTCGGATGGACCACAGGAAGCAATACTTTGGGATGCACATC 4558

Qy       3090 GCACAAAATGAGGCAAATGTGGAATAACAAAAATGACCTTTATCTGGATGATGACCACAT 3149
              ||||||||||||||||||| ||||||||||||||||||||||||||||||||| ||||||
Db       4559 GCACAAAATGAGGCAAATGTGGAATAACAAAAATGACCTTTATCTGGATGATGACCACAT 4618

Qy       3150 TAGTAACCAGCCTTTCTTGAGAGCTCTGTTAGGCCACTGCAGCAATTCTGCAGGAAGCAC 3209
              ||||||||||||||||||| |||||||||||||||||||||||||||   |||||| |
Db       4619 TAGTAACCAGCCTTTCTTGAGAGCTCTGTTAGGCCACTGCAGCAATTCTGCAGGAAGCAC 4678

Qy       3210 CAACTCCTGGGAGTTGAAGAATAAGCAGTTTGAAGCAGCAGTTAAACAAGGAGCATATAG 3269
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4679 CAACTCCTGGGAGTTGAAGAATAAGCAGTTTGAAGCAGCAGTTAAACAAGGAGCATATAG 4738

Qy       3270 CATGGATACAGCCCAGTTTGATGAGCTGATAAGAAACATGAGTCAGCTCATGGAAACCGG 3329
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4739 CATGGATACAGCCCAGTTTGATGAGCTGATAAGAAACATGAGTCAGCTCATGGAAACCGG 4798

Qy       3330 AGAGCTCAGCGATGATCTTGCGTCCCAGCTGATATATCAGCTGGTGGCCGAATTAGCCAA 3389
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4799 AGAGCTCAGCGATGATCTTGCGTCCCAGCTGATATATCAGCTGGTGGCCGAATTAGCCAA 4858

Qy       3390 GGCACAGCCAACACACATGCAGTGGCGGGCATCCAGGAAGAGACACCTCCTGCTGCTCA 3449
              ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db       4859 GGCACAGCCAACACACATGCAGTGGCGGGCATCCAGGAAGAGACACCTCCTGCTGCTCA 4918

Qy       3450 GCTCAGAGGGGAAACATTGATTGTGCCCAAAGCTCTCATGCAAAAAACTCATTCAACCTG 3509
               ||||||||||||||      |||||||||||||| |||||   ||||   |     |
Db       4919 GCTCAGAGGGGAAACAGTTGATTGTGCCCAAAGCTCTCATGCAAAAAACTCAGCAAACCT 4978

Qy       3510 GACATTCAAGCCGAAAGGACTTGTTCTCATGAGGCAAAGCCAACCTCCCTCAATTTCATT 3569
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4979 GACATTCAAGCCGAAAGGACTTGTTCTCATGAGGCAAAGCCAACCTCCCTCAATTTCATT 5038

Qy       3570 TAATAAAACAATAAATTTCAGGATTGGAAATACAGTATACATTACAAAAAGACAGAGGT 3629
              ||||||||||||| |      |||||||||||||||||||||||     |  |
Db       5039 TAATAAAACAATAAATTTCAGGATTGGAAATACAGTATACATTACAAAAAGACAGAGGT 5098

Qy       3630 CATCAATATACTGGTGGACCTTATTACCCCAGTGAGGCCACATATACATGGACCAAGGA 3689
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5099 CATCAATATACTGGTGGACCTTATTACCCCAGTGAGGCCACATATACATGGACCAAGGA 5158

Qy       3690 TTGAACTTGGTTACAGCCTCAGTAAAAAATAATTTTGGATGGAACTGGGAAGATACATAT 3749
               | |||||||| |||  |||||||||||||||||||||||||||||||      |
Db       5159 TTGAACTTGGTTACAGCCTCAGTAAAAAATAATTTTGGATGGAACTGGGAAGATACATAT 5218

Qy       3750 ACAGAATCCTACAAGGAAAGAACAAGGCATATATGAATGTTCTGTAGCTAATCATCTTGG 3809
              ||||||||||||||||||| ||||||||||||||||||||||||||||||||||| |
Db       5219 ACAGAATCCTACAAGGAAAGAACAAGGCATATATGAATGTTCTGTAGCTAATCATCTTGG 5278

Qy       3810 TTCAGATGTGGAAAGTTCTTCTGGTTCTGTAGCAGAGGCACCTTCTATCTGGTCTTGGA 3869
              |||||||||||| |   |||||||||||||||||||||  |          |
Db       5279 TTCAGATGTGGAAAGTTCTTCTGGTTCTGTAGCAGAGGCACCTTCTATCTGGTCTTGGA 5338

Qy       3870 AAGAAATATCACCAAACCAGAGCACAACCATCCTGTCCTCCCTCCTGGACTCATCTGGA 3929
              ||||||||||||||||||  |||||||||||         ||||| |||| |||
Db       5339 AAGAAATATCACCAAACCAGAGCACAACCATCCTGTCCTCCCTCCTGGACTCATCTGGA 5398

Qy       3930 GGCAGCCCTTGGAGCAAACGTGACAATCCGATGTCCTGTAAAAGGTGTCCCTCAGCCTAA 3989
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5399 GGCAGCCCTTGGAGCAAACGTGACAATCCGATGTCCTGTAAAAGGTGTCCCTCAGCCTAA 5458

Qy       3990 TATAACTTGGTTGAAGAGAGGAGGATCTCTGAGTGGCAATGTTTCCTTGCTTTTCAATGG 4049
              ||||||||||||||||||||||||||||||| |   | || || |||| ||||| |||||
Db       5459 TATAACTTGGTTGAAGAGAGGAGGATCTCTGAGTGGCAATGTTTCCTTGCTTTTCAATGG 5518

Qy       4050 ATCCCTGTTGTTGCAGAATGTTTCCCTTGAAAATGAAGGAACCTACGTCTGCATAGCCAC 4109
              |||||||||||||||||| |||||||||||||| | |||||||||||||| ||||
Db       5519 ATCCCTGTTGTTGCAGAATGTTTCCCTTGAAAATGAAGGAACCTACGTCTGCATAGCCAC 5578

Qy       4110 CAATGCTCTTGGAAAGGCAGTGGCAACATCCGTACTCCACTTGCTGGAACGAAGATGGCC 4169
              |||||||||||||||||||| | ||||||||||||  | ||||||||| | ||||
Db       5579 CAATGCTCTTGGAAAGGCAGTGGCAACATCCGTACTCCACTTGCTGGAACGAAGATGGCC 5638

Qy       4170 AGAGAGTAGAATCGTATTTCTGCAAGGACATAAAAAGTACATTCTCAGGCAACTAACAC 4229
              |||||||||||||||||||||||||||||||||||||  |||||||||||||||||||||
Db       5639 AGAGAGTAGAATCGTATTTCTGCAAGGACATAAAAAGTACATTCTCAGGCAACTAACAC 5698

Qy       4230 TAGAACCAACAGCAATGACCCAACAGGAGAACCCCGGCTCAAGAGCCTTTTGGGAGCC 4289
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5699 TAGAACCAACAGCAATGACCCAACAGGAGAACCCCGGCTCAAGAGCCTTTTGGGAGCC 5758

Qy       4290 TGGTAACTGGTCACATTGTTCTGCCACCTGTGGTCATTTGGGAGCCCGCATTCAGAGACC 4349
              ||||||||||||||||| ||||||||||||||||| ||||||||||||||||||||||||
Db       5759 TGGTAACTGGTCACATTGTTCTGCCACCTGTGGTCATTTGGGAGCCCGCATTCAGAGACC 5818

Qy       4350 CCAGTGTGTGATGGCCAATGGGCAGGAAGTGAGTGAGGCCCTGTGTGATCACCTCCAGAA 4409
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5819 CCAGTGTGTGATGGCCAATGGGCAGGAAGTGAGTGAGGCCCTGTGTGATCACCTCCAGAA 5878

Qy       4410 GCCACTGGCTGGGTTTGAGCCCTGTAACATCCGGGACTGCCCAGCGAGGTGGTTCACAAG 4469
              |||||||||||||||| |  ||||||||||||||||||||||||||||||||||||||||
Db       5879 GCCACTGGCTGGGTTTGAGCCCTGTAACATCCGGGACTGCCCAGCGAGGTGGTTCACAAG 5938

Qy       4470 TGTGTGGTCACAGTGCTCGGTGTCTTGCGGTGAAGGATACCACAGTCGGCAGGTGACGTG 4529
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5939 TGTGTGGTCACAGTGCTCGGTGTCTTGCGGTGAAGGATACCACAGTCGGCAGGTGACGTG 5998

Qy       4530 CAAGCGGACAAAAGCCAAGGAACTGGTCAGCCTGCTGCTCCCAACAGCATCCCTCCCAA 4589
              ||| ||||||| ||||||    ||||||||||||||||| |  | ||||||||||||
Db       5999 CAAGCGGACAAAAGCCAAGGAACTGGTCAGCCTGCTGCTCCCAACAGCATCCCTCCCAA 6058

Qy       4590 AGACCGGCCTCTGGGAAGAAAACCATGTTTGGTCATCCATGTGTTCAGTGGGAACCAGG 4649
              ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db       6059 AGACCGGCCTCTGGGAAGAAAACCATGTTTGGTCATCCATGTGTTCAGTGGGAACCAGG 6118

Qy       4650 GAACCGGTGTCCTGGACGTCTCATGGCCCGGCTCCGGACCAGCAGCAGCCCACACAGC 4709
               |||||||| ||| |||  |||||||||| |     |  ||||| |||| |||||||
Db       6119 GAACCGGTGTCCTGGACGTCTCATGGCCCGGCTCCGGACCAGCAGCAGCCCACACAGC 6178

Qy       4710 TTGTCAACACAACAGCTCCGACTCCAACTGTGATGACAGAAAGAGACCCACCTTAAGAAG 4769
              |||||||||||||||||||  ||||||||||||||||||||||||||||| |||||
Db       6179 TTGTCAACACAACAGCTCCGACTCCAACTGTGATGACAGAAAGAGACCCACCTTAAGAAG 6238

Qy       4770 GAACTGCACATCAGGGCCGGTGATGTGTGTTGCCACACAGGCCCTTGGAACCCCTGTAC 4829
              ||||||||||||||||||   |||||||||||||||||| |||||  ||||||||||
Db       6239 GAACTGCACATCAGGGCCGGTGATGTGTGTTGCCACACAGGCCCTTGGAACCCCTGTAC 6298

Qy       4830 AGCAGCCTGTGGCAGGGGTTCCCAGTCTGGAAATCGACTGTATCCACACAAGGAGTTG 4889
              |||||||||||||||||| | ||||||||||| |||||||||||||||||| ||||||||
Db       6299 AGCAGCCTGTGGCAGGGGTTCCCAGTCTGGAAATCGACTGTATCCACACAAGGAGTTG 6358

Qy       4890 CAAACCTGTGGCCAAGAGACACTGTGTACAGAAAAAGAAACCAATTTCCTGGCGGCACTG 4949
              |||||||||| |||||| | ||| |||||||||||| |||||||||||| ||||||||
Db       6359 CAAACCTGTGGCCAAGAGACACTGTGTACAGAAAAAGAAACCAATTTCCTGGCGGCACTG 6418
```

```
Qy   4950 TCTTGGGCCCTCCTGTGATAGAGACTGCACAGACACAACTCACTACTGTATGTTTGTAAA 5009
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6419 TCTTGGGCCCTCCTGTGATAGAGACTGCACAGACACAACTCACTACTGTATGTTTGTAAA 6478

Qy   5010 ACATCTTAATTTGTGTTCTCTAGACCGCTACAAACAAAGGTGCTGCCAGTCATGTCAAGA 5069
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6479 ACATCTTAATTTGTGTTCTCTAGACCGCTACAAACAAAGGTGCTGCCAGTCATGTCAAGA 6538

Qy   5070 GGGATAA 5076
          |||||||
Db   6539 GGGATAA 6545


RESULT 7
AAC74709
ID   AAC74709 standard; cDNA; 5343 BP.
XX
AC   AAC74709;
XX
DT   08-FEB-2001  (first entry)
XX
DE   Human ORFX ORF264 polynucleotide sequence SEQ ID NO:527
XX
KW   Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
KW   vulnerary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
KW   anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KW   immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW   hypotensive; dermatological; immunosuppressive; antiinflammatory;
KW   antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
KW   antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
KW   neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW   cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW   cholesterol ester storage; systemic lupus erythematosus; infection;
KW   severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW   allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW   bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW   thrombosis; contraceptive; ss.
XX
OS   Homo sapiens.
XX
PN   WO200058473-A2.
XX
PD   05-OCT-2000.
XX
PF   31-MAR-2000; 2000WO-US08621.
XX
PR   31-MAR-1999;   99US-0127607.
PR   02-APR-1999;   99US-0127636.
PR   05-APR-1999;   99US-0127728.
PR   30-MAR-2000; 2000US-0540763.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Leach M;
XX
DR   WPI; 2000-602362/57.
DR   P-PSDB; AAB40500.
XX
PT   Novel nucleic acids and peptides derived from open reading frame X,
PT   useful for treating e.g. cancers, proliferative disorders,
PT   neurodegenerative disorders and cardiovascular disease -
XX
PS   Claim 5; Page 690-693; 5507pp; English.
XX
CC   AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
CC   which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC   sequences have activities such as cytostatic; hepatotropic; vulnerary;
CC   antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
CC   osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
CC   immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
CC   antidiabetic; hypotensive; dermatological; immunosuppressive;
CC   antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
CC   antithyroid; and antianaemic. The sequences can be used for determining
CC   the presence of or predisposition to, or preventing or treating
CC   pathological conditions associated with an ORFX-associated disorder. The
CC   nucleic acids can be used to express ORFX proteins in gene therapy
CC   vectors. The proteins and nucleic acids may be used to treat cancers,
CC   proliferative disorders, neurodegenerative disorders, osteoarthritis,
CC   graft vs host disease, cardiovascular disease, diabetes mellitus,
CC   hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
CC   erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
CC   bacterial or fungal infection, malaria, autoimmune disorders, asthma,
CC   allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
CC   nocturnal haemoglobinuria, antiinflammatory disease; to enhance
CC   coagulation; to inhibit thrombosis; and as a contraceptive.
XX
SQ   Sequence 5343 BP; 1496 A; 1137 C; 1304 G; 1404 T; 2 other;

  Query Match             64.8%;  Score 3289.6;  DB 21;  Length 5343;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 3303;  Conservative    0;  Mismatches    4;  Indels    1;  Gaps    1;

Qy   1770 CTGCCCTGTGCTCCTCACATTCACCCACACTCAGACTCACCTCCCCCAGCAACAGTGTGA 1829
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3 CTGCCCTGTGCTCCTCACATTCACCCACACTCAGACTCACCTCCCCCAGCAACAGTGTGA 62

Qy   1830 AGGCCCCAAGCTGCCCACCGAACGGCCCTGCCTCCTGGAAGCATGTGATGAGAGCCCGGC 1889
          ||||||||||||||||||||||||||||||| ||||||||||||| ||||||||||||||
Db     63 AGGCCCCAAGCTGCCCACCGAACGGCCCTGCTTCCTGGAAGCATGCGATGAGAGCCCGGC 122

Qy   1890 CTCCCGAGAGCTAGACATCCCTCTCCCTGAGGACACTGACACGACTTACGACTGGGAGTA 1949
          ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db    123 CTCCCGAGAGCTAGACATCCCTCTCCCTGAGGACACTGACACGCCTTACGACTGGGAGTA 182

Qy   1950 CCCTCCCTTCACCCCTTGCACAGCAACATCCTTCGCAGCCCATCAAGAAGCCATACCACT 2009
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    183 CCCTCCCTTCACCCCTTGCACAGCAACATCCTTCGCAGCCCATCAAGAAGCCATACCACT 242

Qy   2010 CTGCTTACATATCCAGACCCAGCAGACACTCAATCACACCTTGTGTGATATGCTCCACCG 2069
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    243 CTGCTTACATATCCAGACCCAGCAGACACTCAATCACACCTTGTGTGATATGCTCCACCG 302

Qy   2070 TCCTCCAGCCATGAGCCAGGCCTGTAACACAGAGCCCTGTCCCCCCAGCTGGCATGTGGC 2129
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    303 TCCTCCAGCCATGAGCCAGGCCTGTAACACAGAGCCCTGTCCCCCCAGCTGGCATGTGGC 362

Qy   2130 CTCTTGGGGGCCCTGCTCAGCTACCTGTGGAGTTGGAATTCAGACCCGAGATGTGTACTG 2189
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    363 CTCTTGGGGGCCCTGCTCAGCTACCTGTGGAGTTGGAATTCAGACCCGAGATGTGTACTG 422

Qy   2190 CCTGCACCCAGGCGAGACCCCTGCCCCTCCTCACGACTGCCCAGATGAAAAGCCCCATCC 2249
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    423 CCTGCACCCAGGCGAGACCCCTGCCCCTCCTCACGACTGCCCAGATGAAAAGCCCCATCC 482

Qy   2250 TTTACAAGCATGCAATCAGTTTGACTGCCCTCCTGGCTGGCACATTGAAGAATGGCAGCA 2309
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    483 TTTACAAGCATGCAATCAGTTTGACTGCCCTCCTGGCTGGCACATTGAAGAATGGCAGCA 542

Qy   2310 CTGTTCCAGGACTTGTGGCGGGGCAACTCAGAACAGAAGACTCACCTGTCGGCAGCTGCT 2369
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    543 CTGTTCCAGGACTTGTGGCGGGGCAACTCAGAACAGAAGACTCACCTGTCGGCAGCTGCT 602

Qy   2370 AACGGATGGCAGCTTTTTGAATCTCTCAGATGAATTGTGCCAAGGACCCAAGGCATCGTC 2429
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    603 AACGGATGGCAGCTTTTTGAATCTCTCAGATGAATTGTGCCAAGGACCCAAGGCATCGTC 662

Qy   2430 TCACAAGTCCTGTGCCAGGACAGACTGTCCTCCACATTTAGCTGTGGGAGACTGGTCGAA 2489
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    663 TCACAAGTCCTGTGCCAGGACAGACTGTCCTCCACATTTAGCTGTGGGAGACTGGTCGAA 722

Qy   2490 GTGTTCTCTCAGTTGTGTTGTTGGAATCCAGAGAAGAAAGCAGGTGTGTCAAAGGCTGGC 2549
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    723 GTGTTCTCTCAGTTGTGTTGTTGGAATCCAGAGAAGAAAGCAGGTGTGTCAAAGGCTGGC 782

Qy   2550 AGCTAAAGGTGGGGCCATCCCCCCATTGAGATGATGGGTAGGGATGTACCAGGGCTTTGC 2609
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db        783 AGTCAAAGGTGGTGGCATCCTCAGTGAGATGATGTGCAGGGATCTACCAGGGCTCCC 842

Qy       2610 TCTTGTAAGATCTTGCCAGATGCCTGAGTGCAGTAAAATCAAATCAGAGATGAAGACAAA 2669
              |||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        843 TCTTGTAAGATCTTGCCAGATGCCTGAGTGCAGTAAAATCAAATCAGAGATGAAGACAAA 902

Qy       2670 ACTTGGTGAGCAGGGTCCGCAGATCCTCAGTGTCAGAGAGGTCTACATTCAGACAAGGGA 2729
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||| |||
Db        903 ACTTGGTGAGCAGGGTCCGCAGATCCTCAGTGTCAGAGAGGTCTACATTCAGACAAGGGA 962

Qy       2730 AGAGAAGCGTATTAACCTGACCATTGGTAGCAGAGCCTATTTGCTGCCCAACACATCCGT 2789
              ||||||||||||| |||| |||||||||||||||||||||||||||||||||||||||||
Db        963 AGAGAAGCGTATTAACCTGACCATTGGTAGCAGAGCCTATTTGCTGCCCAACACATCCGT 1022

Qy       2790 GATTATTAAGTGCCCCGTGCGACGATTCCAGAAATCTCTGATCCAGTGGGAGAAGGATGG 2849
              |||||||||||| ||||||| |||||||||||||||||||||||||||||| | ||||||
Db       1023 GATTATTAAGTGCCCCGTGCGACGATTCCAGAAATCTCTGATCCAGTGGGAGAAGGATGG 1082

Qy       2850 CCGTTGCCTGCAGAACTCCAAACGGCTTGGCATCACCAAGTCAGGCTCACTAAAAATCCA 2909
              ||||||||||| |||||||||||||||||||||||||||||||||||||||| ||| | |
Db       1083 CCGTTGCCTGCAGAACTCCAAACGGCTTGGCATCACCAAGTCAGGCTCACTAAAAATCCA 1142

Qy       2910 CGGTCTTGCTGCCCCCGACATCGGCGTGTACCGGTGCATTGCAGGCTCTGCACAGGAAAC 2969
              ||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db       1143 CGGTCTTGCTGCCCCCGACATCGGCGTGTACCGGTGCATTGCAGGCTCTGCACAGGAAAC 1202

Qy       2970 AGTTGTGCTCAAGCTCATTGGTACTGACAACCGGCTCATCGCACGCCCAGCCCTCAGGGA 3029
              |||||||||||||||||| ||||||||||||||||||||||||||||| ||||||  |||
Db       1203 AGTTGTGCTCAAGCTCATTGGTACTGACAACCGGCTCATCGCACGCCCAGCCCTCAGGGA 1262

Qy       3030 GCCTATGAGGGAATATCCTGGGATGGACCACAGCGAAGCCAATAGTTTGGGAGTCACATG 3089
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1263 GCCTATGAGGGAATATCCTGGGATGGACCACAGCGAAGCCAATAGTTTGGGAGTCACATG 1322

Qy       3090 GCACAAAATGAGGCAAATGTGGAATAACAAAAATGACCTTTATCTGGATGATGACCACAT 3149
              ||||||||||||||||||   |||||||||||||||||||||||||||||| | || ||
Db       1323 GCACAAAATGAGGCAAATGTGGAATAACAAAAATGACCTTTATCTGGATGATGACCACAT 1382

Qy       3150 TAGTAACCAGCCTTTCTTGAGAGCTCTGTTAGGCCACTGCAGCAATTCTGCAGGAAGCAC 3209
              |||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db       1383 TAGTAACCAGCCTTTCTTGAGAGCTCTGTTAGGCCACTGCAGCAATTCTGCAGGAAGCAC 1442

Qy       3210 CAACTCCTGGGAGTTGAAGAATAAGCAGTTTGAAGCAGCAGTTAAACAAGGAGCATATAG 3269
              |||||||||||||| || ||| ||||||||||||||||||||||||||| || || ||
Db       1443 CAACTCCTGGGAGTTGAAGAATAAGCAGTTTGAAGCAGCAGTTAAACAAGGAGCATATAG 1502

Qy       3270 CATGGATACAGCCCAGTTTGATGAGCTGATAAGAAACATGAGTCAGCTCATGGAAACCGG 3329
              |||||||||||||||| ||||||||||||||||||||||||||||||||||| || | |
Db       1503 CATGGATACAGCCCAGTTTGATGAGCTGATAAGAAACATGAGTCAGCTCATGGAAACCGG 1562

Qy       3330 AGAGGTCAGCGATGATCTTGCGTCCCAGCTGATATATCAGCTGGTGGCCGAATTAGCCAA 3389
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db       1563 AGAGGTCAGCGATGATCTTGCGTCCCAGCTGATATATCAGCTGGTGGCCGAATTAGCCAA 1622

Qy       3390 GGCACAGCCAACACACATGCAGTGGCGGGGCATCCAGGAAGAGACACCTCCTCCTCCTCA 3449
              ||||||||||||||||| |||||||||||||||||||||||||||||||     |   ||
Db       1623 GGCACAGCCAACACACATGCAGTGGCGGGGCATCCAGGAAGAGACACCTCCTCCTCCTCA 1682

Qy       3450 GCTCAGAGGGGAAACAGGGAGTGTGTCCCAAAGCTCGCATGCAAAAAACTCAGGCAAGCT 3509
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1683 GCTCAGAGGGGAAACAGGGAGTGTGTCCCAAAGCTCGCATGCAAAAAACTCAGGCAAGCT 1742

Qy       3510 GACATTCAAGCCGAAAGGACCTGTTCTCATGAGGCAAAGCCAACCTCCCTCAATTTCATT 3569
              |||||||||||||||||||| ||||||||||||| |||||||||||||| |||||||||
Db       1743 GACATTCAAGCCGAAAGGACCTGTTCTCATGAGGCAAAGCCAACCTCCCTCAATTTCATT 1802

Qy       3570 TAATAAAACAATAAATTCCAGGATTGGAAATACAGTATACATTACAAAAAGGACAGAGGT 3629
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1803 TAATAAAACAATAAATTCCAGGATTGGAAATACAGTATACATTACAAAAAGGACAGAGGT 1862

Qy       3630 CATCAATATACTGTGGGACTTTATTACCCCCAGTGAGGCCACATATACATGGACCAAGGA 3689
              ||||| |||||||||||    |||||||||||||||||||| ||||||||| ||||||| |
Db       1863 CATCAATATACTGTGGGACTTTATTACCCCCAGTGAGGCCACATATACATGGACCAAGGA 1922

Qy       3690 TGGAACCTTGTTACAGCCCTCAGTAAAAATAATTTTGGATGGAACTGGGAAGATACAGAT 3749
              ||||||||||||| ||||||||||||||||||||||||||||||||| ||||||||||||
Db       1923 TGGAACCTTGTTACAGCCCTCAGTAAAAATAATTTTGGATGGAACTGGGAAGATACAGAT 1982

Qy       3750 ACAGAATCCTACAAGGAAAGAACAAGGCATATATGAATGTTCTGTAGCTAATCATCTTGG 3809
              |||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db       1983 ACAGAATCCTACAAGGAAAGAACAAGGCATATATGAATGTTCTGTAGCTAATCATCTTGG 2042

Qy       3810 TTCAGATGTGGAAAGTTCTTCTGTGCTGTATGCAGAGGCACCTGTCATCTTGTCTGTTGA 3869
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2043 TTCAGATGTGGAAAGTTCTTCTGTGCTGTATGCAGAGGCACCTGTCATCTTGTCTGTTGA 2102

Qy       3870 AAGAAATATCACCAAACCAGAGCACAACCATCTGTCTGTTGTCGTTGGAGGCATCGTGGA 3929
              |||||||||||||||||||||||||||||||||||||||||| | |||||||||||||||
Db       2103 AAGAAATATCACCAAACCAGAGCACAACCATCTGTCTGTTGTCGTTGGAGGCATCGTGGA 2162

Qy       3930 GGCAGCCCTTGGAGCAAACGTGACAATCCGATGTCCTGTAAAAGGTGTCCCTCAGCCTAA 3989
              ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db       2163 GGCAGCCCTTGGAGCAAACGTGACAATCCGATGTCCTGTAAAAGGTGTCCCTCAGCCTAA 2222

Qy       3990 TATAACTTGGTTGAAGAGAGGAGGATCTCTGATGGGAATGTTTCCTTTGCTTCTCAATGG 4049
              ||||||||||||||||||| |||||||||||||||||||||||||||||| |||||||||
Db       2223 TATAACTTGGTTGAAGAGAGGAGGATCTCTGATGGGAATGTTTCCTTTGCTTCTCAATGG 2282

Qy       4050 ATCCCTGTTGTTGCAGAAGGTTTCCCTTGAAAATGAAGGAACCTACGTCTGCATAGCCAC 4109
              |||||||||||||||||| | |||||||||||||||||||||||||||||||||||||||
Db       2283 ATCCCTGTTGTTGCAGAAGGTTTCCCTTGAAAATGAAGGAACCTACGTCTGCATAGCCAC 2342

Qy       4110 CAATGCTCTTGGAAAGGCAGTGGCAACATCTGTACTCCACTTGCTGGAACGAAGATGGCC 4169
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2343 CAATGCTCTTGGAAAGGCAGTGGCAACATCTGTACTCCACTTGCTGGAACGAAGATGGCC 2402

Qy       4170 AGAGAGTAGAATCGTATTTCTGCAAGGACATAAAAAGTACATTCTTCAGGCAACCAACAC 4229
              |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db       2403 AGAGAGTAGAATCGTATTTCTGCAAGGACATAAAAAGTACATTCTTCAGGCAACCAACAC 2462

Qy       4230 TAGAACCAACAGCAATGACCCAACAGGAGAACCCTCGCCTTAAGAGCCTTCTTGGGAGCC 4289
              |||||||||||||||||||||||||||||||||||||| |||||| ||||||||||||
Db       2463 TAGAACCAACAGCAATGACCCAACAGGAGAACCCTCGCCTTAAGAGCCTTCTTGGGAGCC 2522

Qy       4290 TGGTAACTGGTCACATTGTTCTGCCACCTGTGGTCATTTGGGAGCCCGCATCCAGAGACC 4349
              |||||||||||||||||||| ||||||||||||||||||||| |||||  | |||||||
Db       2523 TGGTAACTGGTCACATTGTTCTGCCACCTGTGGTCATTTGGGAGCCCGCATCCAGAGACC 2582

Qy       4350 CCAGTGTGTGATGGCCAATGGGCAGGAAGTGAGTGAGGCCCTGTGTGATCA-CCTCCAGA 4408
              ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db       2583 CCAGTGTGTGATGGCCAATGGGCAGGAAGTGAGTGAGGCCCTGTGTGATCACCCTCCAGA 2642

Qy       4409 AGCCACTGGCTGGTTTGAGCCCTGTAACATCCGGGACTGCCCCAGGAGGTGGGTTCATAA 4468
              ||||||||||||| ||||| ||||||||||||||||||| |||||||||||||||||||
Db       2643 AGCCACTGGCTGGTTTGAGCCCTGTAACATCCGGGACTGCCCCAGGAGGTGGGTTCACAA 2702

Qy       4469 GTGTGTGGTCACAGTGCTTGTGTCTTGCCGGTGAAGGATACCACAGTCGGCAGGTGACGT 4528
              |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db       2703 GTGTGTGGTCACAGTGCTTGTGTCTTGCCGGTGAAGGATACCACAGTCGGCAGGTGACGT 2762

Qy       4529 GCAAGCGGACAAAAGCCAATGGAACTGTGCAGGTGGTGTCTCCAAGAGCATCTTGCCCTA 4588
              ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db       2763 GCAAGCGGACAAAAGCCAATGGAACTGTGCAGGTGGTGTCTCCAAGAGCATCTTGCCCTA 2822

Qy       4589 AAGACCGGCCTCTGGGAAGAAAACCATGTTTTGGTCATCCATTGGTTCAGTCGGAACCAG 4648
              |||||||| |||||||||||||||||||||| ||||||||||| ||||||||||||||||
Db       2823 AAGACCGGCCTCTGGGAAGAAAACCATGTTTTGGTCATCCATTGGTTCAGTCGGAACCAG 2882

Qy       4649 GGAACCGGTGTCCTGGACGTTGCATGGGCCGGCCTGTGAGGATGCAGCAGCGTCACACAG 4708
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2883 GGAACCGGTGTCCTGGACGTTGCATGGGCCGGCCTGTGAGGATGCAGCAGCGTCACACAG 2942

Qy       4709 CTTGTCAACAAAACAGCTCTGACTCCAACTGGGATGACAGAAAGAGACCTACCTTAAGAA 4768
              ||||||||||||||||||||| ||||||||| | |||||||||||||||||||||||||
Db       2943 CTTGTCAACAAAACAGCTCTGACTCCAACTGGGATGACAGAAAGAGACCTACCTTAAGAA 3002
```

```
Qy       4769 GGAACTGCACATCAGGGGCCTGTGATGTGTGTTGGCACACAGGCCCTTGGAAGCCCTGTA 4828
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3003 GGAACTGCACATCAGGGGCCTGTGATGTGTGTTGGCACACAGGCCCTTGGAAGCCCTGTA 3062

Qy       4829 CAGCAGCCTGTGGCAGGGGTTTCCAGTCTCGGAAAGTCGACTGTATCCACACAAGGAGTT 4888
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3063 CAGCAGCCTGTGGCAGGGGTTTCCAGTCTCGGAAAGTCGACTGTATCCACACAAGGAGTT 3122

Qy       4889 GCAAACCTGTGGCCAAGAGACACTGTGTACAGAAAAAGAAACCAATTTCCTGGCGGCACT 4948
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3123 GCAAACCTGTGGCCAAGAGACACTGTGTACAGAAAAAGAAACCAATTTCCTGGCGGCACT 3182

Qy       4949 GTCTTGGGCCCTCCTGTGATAGAGACTGCACAGACACAACTCACTACTGTATGTTTGTAA 5008
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3183 GTCTTGGGCCCTCCTGTGATAGAGACTGCACAGACACAACTCACTACTGTATGTTTGTAA 3242

Qy       5009 AACATCTTAATTTGTGTTCTCTAGACCGCTACAAACAAAGGTGCTGCCAGTCATGTCAAG 5068
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3243 AACATCTTAATTTGTGTTCTCTAGACCGCTACAAACAAAGGTGCTGCCAGTCATGTCAAG 3302

Qy       5069 AGGGATAA 5076
              ||||||||
Db       3303 AGGGATAA 3310


RESULT 8
AAF87125
ID   AAF87125 standard; DNA; 2912 BP.
XX
AC   AAF87125;
XX
DT   26-MAR-2002  (first entry)
XX
DE   NOV14 coding sequence.
XX
KW   NOV; Cytostatic; contraceptive; antiinflammatory, immunomodulatory;
KW   cardiovascular; casein kinase II phosphorylation site, contraception,
KW   serine/threonine kinase; Peutz-Jeghers syndrome, cellular proliferation,
KW   epidermal growth factor; cell development, apoptosis, cell adhesion;
KW   growth migration; cell structure, motility, cancer; immune disorder;
KW   inflammatory disorder; cellular adhesion disorder, long-QT syndrome,
KW   cardiovascular disease; hypertrophic cardiomyopathy, marfan syndrome;
KW   therapy; NOV1; NOV2; NOV3; NOV4; NOV5, NOV6; NOV7, NOV8, NOV9, NOV10,
KW   NOV11; NOV12; NOV13; NOV14; NOV15, NOV16; ds.
XX
OS   Unidentified.
XX
FH   Key             Location/Qualifiers
FT   CDS             98..2878
FT                   /*tag=  a
FT                   /product  "NOV14"
XX
PN   WO200136638-A2.
XX
PD   25-MAY-2001.
XX
PF   17-NOV-2000; 2000WO-US31543.
XX
PR   19-NOV-1999;    99US-166336P.
PR   29-NOV-1999;    99US-167785P.
PR   08-MAR-2000; 2000US-187844P.
PR   16-NOV-2000; 2000US-0715417.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Lichenstein H,  Vernet C,  Fernandes E;
XX
DR   WPI; 2001-648134/74.
DR   P-PSDB; AAB83376.
XX
PT   Novel human polypeptides and the nucleic acids that encode them useful
PT   for preventing, diagnosing and treating e.g. cancer, inflammation and
PT   immune disorders -
XX
PS   Claim 9; Page 43-45; 141pp; English.
XX
CC   This sequence encodes the NOV14 protein. The invention relates to
CC   the NOV1-NOV16 proteins, and their coding sequences. The proteins have
CC   Cytostatic; contraceptive; antiinflammatory; immunomodulatory; and
CC   cardiovascular activities. The sequences may be used in the prevention,
CC   diagnosis and treatment of diseases associated with inappropriate NOVX
CC   expression. They may be used to treat disorders associated with decreased
CC   expression by rectifying mutations or deletions in a patient's genome
CC   that affect the activity of protein by expressing inactive proteins or to
CC   supplement the patients own production of protein. They are used to
CC   produce NOVX proteins, by inserting the nucleic acid into a cell and
CC   culturing it to express the protein. The DNA may be used as DNA probes in
CC   assays to detect and quantitate the presence of similar DNAs in samples,
CC   and which patients may need restorative therapy. The NOVX protein may
CC   also be used as antigens in the production of antibodies (Abs) against
CC   NOVX and in assays to identify modulators of NOVX expression and
CC   activity. The anti-NOVX Abs and antagonist are used to down regulate
CC   expression and activity. The anti NOVX Abs are used for detecting the
CC   presence of NOVX in samples. Disorders that may be prevented, diagnosed
CC   and/or treated vary depending on the NOVX protein. NOV1, NOV3, NOV5,
CC   NOV7, NOV9-11 and NOV13-16 have casein kinase II phosphorylation sites
CC   characteristic of serine/threonine kinases, and are used to treat
CC   kinase-related disorders (e.g. Peutz Jeghers syndrome, cellular
CC   proliferation and contraception). NOV2-3, NOV6 and NOV8 are homologous to
CC   the epidermal growth factor (EGF)-like super family and are involved in,
CC   e.g. regulation of cell development, apoptosis, cell adhesion, growth
CC   migration, cell structure and motility and protein management, and are
CC   used to treat cancers, inflammatory disorders, immune disorders and
CC   cellular adhesion disorders. NOV6-10 are homologous to EGF like
CC   fibrillin proteins and are used to treat cardiovascular disease e.g.
CC   hypertrophic cardiomyopathy, long QT syndrome and marfan syndrome.
XX
SQ   Sequence 2912 BP; 810 A; 713 C; 754 G; 635 T; 0 other;

  Query Match             56.4%;  Score 2862.8;  DB 23;  Length 2912;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 2867;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy        843 CAGCTTTAACAGCCCCGGCGTCTTTGTCGTAGAAAACACAACAGTGGAATTTCAGACGGG 902
              ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db          1 CAGCTTTAACAGCCCCGGCGTCTTTGTCGTAGAAAACACAACAGTGGAATTTTAGAGGGG 60

Qy        903 CTCCGAGAGGCAAACTTTTAAGATTCCAGGACCTCTGATGGCTGATTTCATCTTCAAGAC 962
              |||||||||||||||||||||||||||||| ||| |||||||||||||||||||||||||
Db         61 CTCCGAGAGGCAAACTTTTAAGATTCCAGGCCCTTTGATGGCTGATTTCATCTTCAAGAC 120

Qy        963 CAGGTACACTGCAGCCAAAGACAGCGTGGTTCAGTTCTTCTTTACCAGCCCATCAGTCA 1022
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CAGGTACACTGCAGCCAAAGACAGCGTGGTTCAGTTCTTCTTTACCAGCCCATCAGTCA 180

Qy       1023 TCAGTGGAGACAAACTGACTTCTTTCCCTGCACTGTGACGTGTGGAGGAGGTTATCAGCT 1082
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TCAGTGGAGACAAACTGACTTCTTTCCCTGCACTGTGACGTGTGGAGGAGGTTATCAGCT 240

Qy       1083 CAATTCTGCTGAATGTGTGGATATCCGCTTGAAGAGGGTAGTTCCTGACCATTATTGTCA 1142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CAATTCTGCTGAATGTGTGGATATCCGCTTGAAGAGGGTAGTTCCTGACCATTATTGTCA 300

Qy       1143 CTACTACCCTGAAAATGTAAAACCAAAACCAAAACTGAAGGAATGCACCATGGATCCCTC 1202
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CTACTACCCTGAAAATGTAAAACCAAAACCAAAACTGAAGGAATGCAGCATGGATCCCTG 360

Qy       1203 CCCATCAAGTGATGGATTTAAAGAGATAATGCCCTATGACCACTTCCAACCTCTTCCTCG 1262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCCATCAAGTGATGGATTTAAAGAGATAATGCCCTATGACCACTTCCAACCTCTTCCTCG 420

Qy       1263 CTGGGAACATAATCCTTGGACTGCATGTTCCGTGTCCTGTGGAGGAGGGATTCAGAGACG 1322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTGGGAACATAATCCTTGGACTGCATGTTCCGTGTCCTGTGGAGGAGGGATTCAGAGACG 480

Qy       1323 GAGCTTTGTGTGTGTAGAGGAATCCATGCATGGAGAGATATTGCAGGTGGAAGAATGGAA 1382
```

```
             ||||||||||  ||    |  ||||||||||||||||  ||  ||||| |    |
Db      481 CAGCTTTGTTTCGTAGATGAATCATGCATGGAGAGATATTGCACCTGGAAGAACTGAA 540

Qy     1383 GTGCATGTACGCACCCAAACCAAGGTTATGCAAACTTGTAATCTGTTTGATTGCCCCAA 1442
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      541 GTGCATGTACGCACCCAAACCAAGGTTATGCAAACTTGTAATCTGTTTGATTGCCCCAA 600

Qy     1443 GTGGATTGCCATGGAGTGGTTCTAGTGCACAGTGACTTGTGGCCGAGGGTTACGGTACCG 1502
            |||||||||||||||||| | |||||||||||||||||||||| |||||||||||||  |
Db      601 GTGGATTGCCATGGAGTGGTTCTAGTGCACAGTGACTTGTGGCCGAGGGTTACGGTACCG 660

Qy     1503 GGTTGTTCTGTGTATTAACTACCCCGGAGAGCATGTTGGGGCTGCAATCCACAATCCAA 1562
            ||||||||||||||||||    ||||||||||||||||||||| |||| |  | |    |
Db      661 GGTTGTTCTGTGTATTAACCACCCGGGAGAGCATGTTGGGGCTGCAATCCACAATCCAA 720

Qy     1563 GTTACACATCAAAGAAGAAGCCGCTCATTCCCATCCCGCTGTTATAAACCAAAACAAAAAAG 1622
            ||||||||||||||||| || ||||||||||||||||||||||||||||| |||    |
Db      721 GTTACACATCAAAGAAGAATGCGTCATTCCCATCCCGTGTTATAAACCAAAAGAAAAAAG 780

Qy     1623 TCCAGTGGAAGCAAAATTGCCTTGGCTGAAACAAGCACAAGAACTAGAAGAGACCAGAAT 1682
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      781 TCCAGTGGAAGCAAAATTGCCTTGGCTGAAACAAGCACAAGAACTAGAAGAGACCAGAAT 840

Qy     1683 AGCAACAGAAGAACCAACTTTCATTCCAGAACCCTGGTCAGCCTGCAGTACCACGTGTGG 1742
            |||||||||||||||||| ||||||||||||||||||||||||||||||||||| ||||
Db      841 AGCAACAGAAGAACCAACGGTCATTCCAGAACCCTGGTCAGCCTGCAGGACCACGTGTGG 900

Qy     1743 GCCAGGTGTGCAGGTCGCCAGGTGAAGTGCCGTGTGCTCCTCACATTCACGCAGACTGA 1802
            ||| ||||||||||||||  | |||||||||||||||||||||||||||| |  |||| ||
Db      901 GCCGGTGTGCAGGTCCGCGAGGTGAAGTGCCGTGTGCTCCTCACATTCACGCAGACTGA 960

Qy     1803 GACTGAGCTGCCCGAGGAAGAGGTGGAAGGCCCCAAGCTGCCCACCGAACGGCCCTGCCC 1862
            |||||||||||||| |   ||||||||||||||||||||||||||||||||| ||||||
Db      961 GACTGAGCTGCCCGAGGAAGAGGTGGAAGGCCCCAAGCTGCCCACCGAACGGCCCTGCCT 1020

Qy     1863 CCTGGAAGCATGTGATGAGAGCCCGGCCTCCCGAGAGCTAGACATCCCTCTCCCTGAGGA 1922
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1021 CCTGGAAGCATGTGATGAGAGCCCGGCCTCCCGAGAGCTAGACATCCCTCTCCCTGAGGA 1080

Qy     1923 CAGTGAGACGACTTACGACTGCGAGTACGCTGGGTTCACCCCTTGCACAGCAACATGCTT 1982
            |||||||||||||||||||| ||||||||||||||||||||||||| |  |||  |   |
Db     1081 CAGTGAGACGACTTACGACTCCGAGTACGCTGGGTTCACCCCTTGCACAGCAACATGCTT 1140

Qy     1983 GGGAGGCCATCAAGAAGCTATACACAGGGTGCTTACATATCCAGACCCACACAGACAGCAA 2042
            |||||||||||||||||||| ||  |||||||||||||||||||||||||| |  |   |
Db     1141 GGGAGGCCATCAAGAAGCTACACACATGGTGCTTACATATCCAGACCCACACAGACAGCAA 1200

Qy     2043 TGACACCTTGGTTGATATATGGCCTACCGTCCTCCAGCCATGAGCCAGGCCTGAACACAGA 2102
            ||||| ||||  |||||||||||||||||||||||||||||||||||||||||| |  | |
Db     1201 TGACAGCTTGTGTGATATATGGCCTACCGTCCTCCAGCCATGAGCCAGGCCTGAACACAGA 1260

Qy     2103 GCCCTGTCCCCCAGGTGGCATGTGGGCTCTTGGGGGCCCTGCTCAGCTACCTGTGGAGT 2162
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1261 GCCCTGTCCCCCAGGTGGCATGTGGGCTCTTGGGGGCCCTGCTCAGCTACCTGTGGAGT 1320

Qy     2163 TGGAATTCAGACCCGAGATGGGTACTGCCTGGACCCAGGGGAGACCCTGGCCCTCCCTGA 2222
            ||||||||||||||||| |||||||||||||||||||||||||||| ||||||||| |||
Db     1321 TGGAATTCAGACCCGAGATGGGTACTGCCTGGACCCAGGGGAGACCCCGGCCCTCCCTGA 1380

Qy     2223 GGAGTGCCGAGATGAAAACCCCATGCTTTACAAGCATGCAATGCAGCTGGACTGCCCCTG 2282
            |||||||||||||||     ||||||||||||||||||||||||||||| |  |     |
Db     1381 GGAGTGCCGAGATGAAAACCCCATGCTTTACAAGCATGCAATGCAGCTGGACTGCCCCTG 1440

Qy     2283 TGCCTGGCACATTGAAGAATGGCAGCAGTGTTCCAGGACTTGTGCCGGGGGAACTGACAA 2342
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1441 TGCCTGGCACATTGAAGAATGGCAGCAGTGTTCCAGGACTTGTGCCGGGGGAACTGACAA 1500

Qy     2343 CAGAAGAGTCACCTGTCGGCAGCTGCTAACGGATGGCAGCTTTTGAATCTCTCAGAGGA 2402
            |||||||||||||||||||  ||||||||||||||||||||||||||||||||||| ||
Db     1501 CAGAAGAGTCACCTGTCGGCAGCTGCTAACGGATGGCAGCTTTTGAATCTCTCAGAGGA 1560

Qy     2403 ATTGTGCCAAGGACCCAAGGCATGTCTCACAAGTCCTGGCCAGGACAGACTGGTCCCCC 2462
            ||||||||||||||     |||||||||||| |||||||||| | |   ||  |
Db     1561 ATTGTGCCAAGGACCCAAGGCATGTCTCACAAGTCCTGGCCAGGACAGACTGGTCCCCC 1620

Qy     2463 ACATTTAGCTGTGGGAGACTGGTCGAAGTGTTCTGTCACGTGCGGTCTGGGAATCCAGAG 2522
            ||||||||||||||||| | |||||||||||||||||||||||||||||||| |||||||
Db     1621 ACATTTAGCTGTGGGAGACTGGTCGAAGTGTTCTGTCACGTGCGGTCTGGGAATCCAGAG 1680

Qy     2523 AAGAAAGCAGGTGTGTCAAAGGCTGGCAGCCAAAGGTCGGCGCATCCCCCTCAGTGAGAT 2582
            ||||||||||||||||||| | |||||||||||||||||||||||||||||| |||||||
Db     1681 AAGAAAGCAGGTGTGTCAAAGGCTGGCAGCCAAAGGTCGGCGCATCCCCCTCAGTGAGAT 1740

Qy     2583 GATGTGCAGGGATCTACCAGGGTTCCCTCTTGTAAGATCTTGCCAGATGCCTGAGTGCAG 2642
            |||||||||||||||||| | |||||||||||||||||||||||||||||||| ||||||
Db     1741 GATGTGCAGGGATCTACCAGGGTTCCCTCTTGTAAGATCTTGCCAGATGCCTGAGTGCAG 1800

Qy     2643 TAAAATCAAATCAGAGATGAAGACAAAACTTGGTGAGCAGGGTCCGCAGATCCTCAGTGT 2702
            |||||||||||||||||| | ||||||||||||||||||||||||||||||| |||||||
Db     1801 TAAAATCAAATCAGAGATGAAGACAAAACTTGGTGAGCAGGGTCCGCAGATCCTCAGTGT 1860

Qy     2703 CCAGAGAGTCTACATTCAGACAAGGGAAGAGAAGCGTATTAACCTGACCATTGGTAGCAG 2762
            |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db     1861 CCAGAGAGTCTACATTCAGACAAGGGAAGAGAAGCGTATTAACCTGACCATTGGTAGCAG 1920

Qy     2763 AGCCTATTTGCTGCCCAACACATCCGTGATTATTAAGTGCCCCGTGCGACGATTCCAGAA 2822
            ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db     1921 AGCCTATTTGCTGCCCAACACATCCGTGATTATTAAGTGCCCCGTGCGACGATTCCAGAA 1980

Qy     2823 ATCTCTGATCCAGTGGGAGAAGGATGGCCGTTGCCTGCAGAACTCCAAACGGCTTGGCAT 2882
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1981 ATCTCTGATCCAGTGGGAGAAGGATGGCCGTTGCCTGCAGAACTCCAAACGGCTTGGCAT 2040

Qy     2883 CACCAAGTCAGGCTCACTAAAAATCCACGGTCTTGCTGCCCCGACATCGGCGTGTACCG 2942
            |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db     2041 CACCAAGTCAGGCTCACTAAAAATCCACGGTCTTGCTGCCCCGACATCGGCGTGTACCG 2100

Qy     2943 GTGCCATTGCAGGCTCTGCACAGGAAACAGTTGTGCTCAAGCTCATTGGTACTGACAACCG 3002
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2101 GTGCCATTGCAGGCTCTGCACAGGAAACAGTTGTGCTCAAGCTCATTGGTACTGACAACCG 2160

Qy     3003 GCTCATCGCACGCCCAGCCCTCAGGGAGCCTATGAGGGAATATCCTGGGATGGACCACAG 3062
            |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db     2161 GCTCATCGCACGCCCAGCCCTCAGGGAGCCTATGAGGGAATATCCTGGGATGGACCACAG 2220

Qy     3063 CGAAGCCAATAGTTTGGGAGTCACATGGCACAAAATGACCGAAATGGGAATAACAAAAA 3122
            |||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db     2221 CGAAGCCAATAGTTTGGGAGTCACATGGCACAAAATGACCGAAATGGGAATAACAAAAA 2280

Qy     3123 TGACCTTTATCTGGATGATGACCACATTAGTAACCAGCCTTTCTTGAGAGCTCTGTTAGG 3182
            ||||||||||||||||||||||||||||||| ||||||||||||||||||||| ||||||
Db     2281 TGACCTTTATCTGGATGATGACCACATTAGTAACCAGCCTTTCTTGAGAGCTCTGTTAGG 2340

Qy     3183 CCACTGCAGCAATTCTGCAGGAAGCACCAACTCCTGGGAGTTGAAGAATAAGCAGTTTGA 3242
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2341 CCACTGCAGCAATTCTGCAGGAAGCACCAACTCCTGGGAGTTGAAGAATAAGCAGTTTGA 2400

Qy     3243 AGCAGCAGTTAAACAAGGAGCATATAGCATGGATACAGCCCAGTTTGATGAGCTGATAAG 3302
            | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2401 AGCAGCAGTTAAACAAGGAGCATATAGCATGGATACAGCCCAGTTTGATGAGCTGATAAG 2460

Qy     3303 AAACATGAGTCAGCTCATGGAAACCGGAGAGGTCAGCGATGATCTTGCTTCCCAGCTGAT 3362
            |||||||||||||||||     |||||||||||| |||||||||||||||||||||||||
Db     2461 AAACATGAGTCAGCTCATGGAAACCGGAGAGGTCAGCGATGATCTTGCTTCCCAGCTGAT 2520

Qy     3363 ATATCAGCTGGTGGCCGAATTAGCCAAGGCACAGCCAACACACATGCAGTGGCGGGGCAT 3422
            ||||||||||||||||| | |||||||||||||||      ||||||||||||||||||
Db     2521 ATATCAGCTGGTGGCCGAATTAGCCAAGGCACAGCCAACACACATGCAGTGGCGGGGCAT 2580

Qy     3423 CCAGGAAGAGACACCTCCCGCTGCTCAGCTCAGAGGGGAAACAGGGAGTGTGTCCCAAAG 3482
            ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db     2581 CCAGGAAGAGACACCTCCCGCTGCTCAGCTCAGAGGGGAAACAGGGAGTGTGTCCCAAAG 2640

Qy     3483 CTCGGCATGCAAAAAACTCAGGCAAGCGACATCCAAGCCCAAAGGACCTGTCCTCATGAG 3542
            ||||||||||||||||||| | | |||| | | | | |||||||||||||||||||||||
Db     2641 CTCGGCATGCAAAAAACTCAGGCAAGCGACATCCAAGCCCAAAGGACCTGTCCTCATGAG 2700
```

```
Qy     3543 GCAAAGCCAACCTCCCTCAATTTCATTTAATAAAACAATAAATTCCAGGATTGGAAATAC 3602
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2701 GCAAAGCCAACCTCCCTCAATTTCATTTAATAAAACAATAAATTCCAGGATTGGAAATAC 2760

Qy     3603 AGTATACATTACAAAAAGGACACAGGTCATCAATATACTGTGTGACCTTATTACCCCCAG 3662
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2761 AGTATACATTACAAAAAGGACACAGGTCATCAATATACTGTGTGACCTTATTACCCCCAG 2820

Qy     3663 TGAGGCCACATATACATGGACCAAGGATGGAACCTTGTTACAGCCCTCAGTAAA 3716
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2821 TGAGGCCACATATACATGGACCAAGGATGGAACCTTGTTACAGCCCTCAGTAAA 2874


RESULT 9
AAD14369
ID   AAD14369 standard; cDNA; 2538 BP.
XX
AC   AAD14369;
XX
DT   01-NOV-2001  (first entry)
XX
DE   Novel human protein (NHP) cDNA #6.
XX
KW   Novel human protein, NHP, thrombospondin, gene therapy, cancer,
KW   cytostatic; antisense therapy; angiogenesis; biological disorder; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..2538
FT                   /*tag=  a
FT                   /product= "Novel human protein"
XX
PN   WO200161011-A2.
XX
PD   23-AUG-2001.
XX
PF   15-FEB-2001; 2001WO-US05290.
XX
PR   17-FEB 2000; 2000US 0183282.
XX
PA   (LEXI-) LEXICON GENETICS INC.
XX
PI   Donoho G,  Scoville J,  Turner CA,  Friedrich G,  Zambrowicz B,
PI   Sands AT;
XX
DR   WPI; 2001-514776/56.
DR   P-PSDB; AAE07868.
XX
PT   An isolated nucleic acid encoding a thrombospondin useful as a
PT   hybridization probe and gene therapy treatments of cancer -
XX
PS   Disclosure; Page 42-43; 56pp; English.
XX
CC   The invention relates to novel human polynucleotides encoding proteins
CC   that share sequence similarity with animal proteins having thrombospondin
CC   repeats. NHPs are expressed in human cell lines, pituitary, lymph nodes,
CC   prostrate, testis, adrenal gland, uterus, foetal kidney, foetal lung and
CC   gene trapped human cells. NHPs are thrombospondins useful for treating
CC   biological disorders involving angiogenesis, cancer and development and
CC   also in pharmacogenomic applications. NHPs are useful as a hybridisation
CC   probe for screening libraries, assessing gene expression patterns and
CC   also in gene therapy. Proteins having thrombospondin repeats act as
CC   receptors, secreted extracellular matrix proteins and proteases. The
CC   present sequence is a cDNA encoding novel human protein (NHP).
XX
SQ   Sequence 2538 BP; 670 A; 627 C; 678 G; 563 T; 0 other;

  Query Match             49.1%;  Score 2490;  DB 22;  Length 2538;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 2490;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy        1 ATGGCTTCCTGGACGAGCCCCTGGTGGGTGCTGATAGGGATGGTCTTCATGCACTCTCCC 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1 ATGGCTTCCTGGACGAGCCCCTGGTGGGTGCTGATAGGGATGGTCTTCATGCACTCTCCC 60

Qy       61 CTCCCGCAGACCACAGCTGAGAAATCTCCTGGAGCCTATTTCCTTCCCGAGTTTGCACTT 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61 CTCCCGCAGACCACAGCTGAGAAATCTCCTGGAGCCTATTTCCTTCCCGAGTTTGCACTT 120

Qy      121 TCTCCTCAGGGAAGTTTTCTGGAAGACACAACAGGGGAGCAGTTCCTCACTTATCGCTAT 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      121 TCTCCTCAGGGAAGTTTTCTGGAAGACACAACAGGGGAGCAGTTCCTCACTTATCGCTAT 180

Qy      181 GATGACCAGACCTCAAGAAACACTCGTTCAGATGAAGACAAAGATGGCAACTGGGATGCT 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      181 GATGACCAGACCTCAAGAAACACTCGTTCAGATGAAGACAAAGATGGCAACTGGGATGCT 240

Qy      241 TGGGGCGACTGGAGTGACTGCTCCCGGACCTGTGGGGGAGGAGCATCATATTCTCTGCGG 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241 TGGGGCGACTGGAGTGACTGCTCCCGGACCTGTGGGGGAGGAGCATCATATTCTCTGCGG 300

Qy      301 AGATGTTTGACTGGAAGGAATTGTGAAGGGCAGAACATTCGGTACAAGACATGCAGCAAT 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      301 AGATGTTTGACTGGAAGGAATTGTGAAGGGCAGAACATTCGGTACAAGACATGCAGCAAT 360

Qy      361 CATGACTGCCCTCCAGATGCAGAAGATTTCAGAGCCCAGCAGTGCTCAGCCTACAATGAT 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      361 CATGACTGCCCTCCAGATGCAGAAGATTTCAGAGCCCAGCAGTGCTCAGCCTACAATGAT 420

Qy      421 GTCCAGTATCAGGGGCATTACTATGAATGGCTTCCACGATATAATGATCCTGCTGCCCCG 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      421 GTCCAGTATCAGGGGCATTACTATGAATGGCTTCCACGATATAATGATCCTGCTGCCCCG 480

Qy      481 TGTGCACTCAAGTGTCATGCACAAGGACAAAACTTGGTGGTGGAGCTGGCACCTAAGGTA 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      481 TGTGCACTCAAGTGTCATGCACAAGGACAAAACTTGGTGGTGGAGCTGGCACCTAAGGTA 540

Qy      541 CTGGATGGAACTCGTTGCAACACGGACTCCTTGGACATGTGTATCAGTGGCATCTGTCAG 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      541 CTGGATGGAACTCGTTGCAACACGGACTCCTTGGACATGTGTATCAGTGGCATCTGTCAG 600

Qy      601 GCAGTGGGCTGCGATCGGCAACTGGGAAGCAATGCCAAGGAGGACAACTGTGGAGTCTGT 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      601 GCAGTGGGCTGCGATCGGCAACTGGGAAGCAATGCCAAGGAGGACAACTGTGGAGTCTGT 660

Qy      661 GCCGGCGATGGCTCCACCTGCAGGCTTGTACGGGGACAATCAAAGTCACACGTTTCTCCT 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      661 GCCGGCGATGGCTCCACCTGCAGGCTTGTACGGGGACAATCAAAGTCACACGTTTCTCCT 720

Qy      721 GAAAAAAGAGAAGAAAATGTAATTGCTGTTCCTTTGGGAAGTCGAAGTGTGAGAATTACA 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      721 GAAAAAAGAGAAGAAAATGTAATTGCTGTTCCTTTGGGAAGTCGAAGTGTGAGAATTACA 780

Qy      781 GTGAAAGGACCTGCCCACCTCTTTATTGAATCAAAAACACTTCAAGGAAGCAAAGGAGAA 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      781 GTGAAAGGACCTGCCCACCTCTTTATTGAATCAAAAACACTTCAAGGAAGCAAAGGAGAA 840

Qy      841 CACAGCTTTAACAGCCCCGGCGTCTTTGTCGTAGAAAACACAACAGTGGAATTTCAGAGG 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      841 CACAGCTTTAACAGCCCCGGCGTCTTTGTCGTAGAAAACACAACAGTGGAATTTCAGAGG 900

Qy      901 GGCTCCGAGAGGCAAACTTTTAAGATTCCAGGACCTCTGATGGCTGATTTCATCTTCAAG 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      901 GGCTCCGAGAGGCAAACTTTTAAGATTCCAGGACCTCTGATGGCTGATTTCATCTTCAAG 960

Qy      961 ACCAGGTACACTGCAGCCAAAGACAGCGTGGTTCAGTTCTTCTTTTACCAGCCCATCAGT 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      961 ACCAGGTACACTGCAGCCAAAGACAGCGTGGTTCAGTTCTTCTTTTACCAGCCCATCAGT 1020

Qy     1021 CATCAGTGGAGACAAACTGACTTCTTTCCCTGCACTGTGACGTGTGGAGGAGGTTATCAG 1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1021 CATCAGTGGAGACAAACTGACTTCTTTCCCTGCACTGTGACGTGTGGAGGAGGTTATCAG 1080

Qy     1081 CTCAATTCTGCTGAATGTGTGGATATCCGCTTGAAGAGGGTAGTTCCTGACCATTATTGT 1140
```

```
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CTCAATTCTGCTGAATTGCTGATATCCCTTGAAGAGGTAGTTCCTGACCATTATTTT 1140

Qy       1141 CACTACTACCCTGAAAATGTAAAACCAAAACCAAAACTGAAGGAATGCAGCATGGATCCC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CACTACTACCCTGAAAATGTAAAACCAAAACCAAAACTGAAGGAATGCAGCATGGATCCC 1200

Qy       1201 TGCCCATCAAGTGATTGATTTAAAGAGATAATGCCCTATGACCACTTCCAACCTCTTCCT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TGCCCATCAAGTGATTGATTTAAAGAGATAATGCCCTATGACCACTTCCAACCTCTTCCT 1260

Qy       1261 CCCTGGGAACATAATCCTTTGGACTGCATGTTCCGGGTCCTGTGGAGGAGTGATTCAGATA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CCCTGGGAACATAATCCTTTGGACTGCATGTTCCGGGTCCTGTGGAGGAGTGATTCAGATA 1320

Qy       1321 CGGAGCTTTGTGTGTGTAGAGGAATCCATGCATGGAGAGATATTGCAGGTGGAAGAATGC 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CGGAGCTTTGTGTGTGTAGAGGAATCCATGCATGGAGAGATATTGCAGGTGGAAGAATGC 1380

Qy       1381 AAGTGCATGTAGGCACTCAAACCCAAGGTTATGCAAACTTGTAATCTGTTTGATTGGCCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 AAGTGCATGTAGGCACTCAAACCCAAGGTTATGCAAACTTGTAATCTGTTTGATTGGCCC 1440

Qy       1441 AAGTGGATTGCCATGGAGTTGTCTCAGTGCACAGTGACTTGTGGCCGAGGTTACCGTACT 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 AAGTGGATTGCCATGGAGTTGTCTCAGTGCACAGTGACTTGTGGCCGAGGTTACCGTACT 1500

Qy       1501 CGGGTTCTTCTGTGTATTAACTACCGGGAGAGCATGTTGGGGGCTGCAATCCACAACTCT 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 CGGGTTCTTCTGTGTATTAACTACCGGGAGAGCATGTTGGGGGCTGCAATCCACAACTCT 1560

Qy       1561 AAGTTACACATCAAAGAAGAATCCTCTCACCTCCACCTCCTTTTATAAACCAAAACAAAAA 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 AAGTTACACATCAAAGAAGAATCCTCTCACCTCCACCTCCTTTTATAAACCAAAACAAAAA 1620

Qy       1621 AGTCCAGTGGAAGCAAAATCCCTTGGCTGAAACAAGCACAAGAACTAGAAGAGACCAGA 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 AGTCCAGTGGAAGCAAAATCCCTTGGCTGAAACAAGCACAAGAACTAGAAGAGACCAGA 1680

Qy       1681 ATAGCAACAGAAGAACCAACTTTCATTCCAGAACCCTGGTCAGCCTGCAGTACCACTTGT 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 ATAGCAACAGAAGAACCAACTTTCATTCCAGAACCCTGGTCAGCCTGCAGTACCACTTGT 1740

Qy       1741 GGGCCAGGTGTGCAGGTCCGGAGGTGAAGTGCCGTGTGCTCCTACATTCACGCAGACT 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 GGGCCAGGTGTGCAGGTCCGGAGGTGAAGTGCCGTGTGCTCCTACATTCACGCAGACT 1800

Qy       1801 GAGACTGAGCTGCCCGAGGAAGAGTGTGAAGGCCCCAAGCTGCCCACCGAACGGCCCTGC 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 GAGACTGAGCTGCCCGAGGAAGAGTGTGAAGGCCCCAAGCTGCCCACCGAACGGCCCTGC 1860

Qy       1861 CTCCTGGAAGCATGTGATGAGAGCCCGGCCTCCCGAGAGCTAGACATCCCTCTCCCTGAG 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 CTCCTGGAAGCATGTGATGAGAGCCCGGCCTCCCGAGAGCTAGACATCCCTCTCCCTGAG 1920

Qy       1921 GACAGTGAGACGACTTACGACTGGGAGTACGCTGGGTTCACCCCTTGCACAGCAACATGC 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 GACAGTGAGACGACTTACGACTGGGAGTACGCTGGGTTCACCCCTTGCACAGCAACATGC 1980

Qy       1981 TTGGGAGGCCATCAAGAAGCCATAGCAGTGTGCTTACATATCCAGACCCAGCAGACAGTC 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 TTGGGAGGCCATCAAGAAGCCATAGCAGTGTGCTTACATATCCAGACCCAGCAGACAGTC 2040

Qy       2041 AATGACAGCTTGTGTGATATCGTCCACCGTCCTCCAGCCATGAGCCAGGCCTGTAACACA 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 AATGACAGCTTGTGTGATATCGTCCACCGTCCTCCAGCCATGAGCCAGGCCTGTAACACA 2100

Qy       2101 GAGCCCTGTCCCCCCAGGTTGGATGTGGGCTCTTGGGGGCCCTGCTCAGCTACCTGTGGA 2160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 GAGCCCTGTCCCCCCAGGTTGGATGTGGGCTCTTGGGGGCCCTGCTCAGCTACCTGTGGA 2160

Qy       2161 GTTGGAATTCAGACCCGAGATGTGTACTGCCTGCACCCAGGGGAGACCCCTGCCCCTCCT 2220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 GTTGGAATTCAGACCCGAGATGTGTACTGCCTGCACCCAGGGGAGACCCCTGCCCCTCCT 2220

Qy       2221 GAGGAGTGCCGAGATGAAAAGCCCCATGCTTTACAAGCATGCAATCAGTTTGACTGCCCT 2280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 GAGGAGTGCCGAGATGAAAAGCCCCATGCTTTACAAGCATGCAATCAGTTTGACTGCCCT 2280

Qy       2281 CCTGGCTGGCACATTGAAGAATGGCAGCAGTGTTCCAGGACTTGTGGCGGGGGAACTCAG 2340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 CCTGGCTGGCACATTGAAGAATGGCAGCAGTGTTCCAGGACTTGTGGCGGGGGAACTCAG 2340

Qy       2341 AACAGAAGAGTCACCTGTCCGGCAGCTGCTAACGGATGGCAGCTTTTGAATCTCTCAGAT 2400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 AACAGAAGAGTCACCTGTCCGGCAGCTGCTAACGGATGGCAGCTTTTGAATCTCTCAGAT 2400

Qy       2401 GAATTGTGCCAAGGACCCAAGGCATCGTCTCACAAGTCCTGTGCCAGGACAGACTGTCCT 2460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 GAATTGTGCCAAGGACCCAAGGCATCGTCTCACAAGTCCTGTGCCAGGACAGACTGTCCT 2460

Qy       2461 CCACATTTAGCTGTGGGAGACTGGTCGAAG 2490
              ||||||||||||||||||||||||||||||
Db       2461 CCACATTTAGCTGTGGGAGACTGGTCGAAG 2490


RESULT 10
AAD14370
ID   AAD14370 standard; cDNA; 2316 BP.
XX
AC   AAD14370;
XX
DT   01-NOV-2001  (first entry)
XX
DE   Novel human protein (NHP) cDNA #7.
XX
KW   Novel human protein; NHP; thrombospondin; gene therapy; cancer;
KW   cytostatic; antisense therapy; angiogenesis; biological disorder; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..2316
FT                   /*tag=  a
FT                   /product= "Novel human protein"
XX
PN   WO200161011-A2.
XX
PD   23-AUG-2001.
XX
PF   15-FEB-2001; 2001WO-US05290.
XX
PR   17-FEB-2000; 2000US-0183282.
XX
PA   (LEXI-) LEXICON GENETICS INC.
XX
PI   Donoho G,  Scoville J,  Turner CA,  Friedrich G,  Zambrowicz B;
PI   Sands AT;
XX
DR   WPI; 2001-514776/56.
DR   P-PSDB; AAE07869.
XX
PT   An isolated nucleic acid encoding a thrombospondin, useful as a
PT   hybridization probe and gene therapy treatments of cancer.
XX
PS   Disclosure; Page 45-46; 56pp; English.
XX
CC   The invention relates to novel human polynucleotides encoding proteins
CC   that share sequence similarity with animal proteins having thrombospondin
CC   repeats. NHPs are expressed in human cell lines, pituitary, lymph nodes,
CC   prostate, testis, adrenal gland, uterus, foetal kidney, foetal lung and
CC   gene trapped human cells. NHPs are thrombospondins useful for treating
CC   biological disorders involving angiogenesis, cancer and development and
CC   also in pharmacogenomic applications. NHPs are useful as a hybridisation
CC   probe for screening libraries, assessing gene expression patterns and
CC   also in gene therapy. Proteins having thrombospondin repeats act as
CC   receptors, secreted extracellular matrix proteins and proteases. The
CC   present sequence is a cDNA encoding novel human protein (NHP).
XX
SQ   Sequence 2316 BP; 617 A; 575 C; 616 G; 508 T; 0 other;

  Query Match             42.9%;  Score 2180;  DB 22;  Length 2316;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 2180;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        311 CTGGAAGGAATTGTGAAGGGCAGAACATTCGGTACAAGACATGCAGCAATCATGACTGCC 370
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         89 CTGGAAGGAATTGTGAAGGGCAGAACATTCGGTACAAGACATGCAGCAATCATGACTGCC 148

Qy        371 CTCCAGATGCAGAAGATTTCAGAGCCCAGCAGTGCTCAGCCTACAATGATGTCCAGTATC 430
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        149 CTCCAGATGCAGAAGATTTCAGAGCCCAGCAGTGCTCAGCCTACAATGATGTCCAGTATC 208

Qy        431 AGGGGCATTACTATGAATGGCTTCCACGATATAATGATCCTGCTGCCCCGTGTGCACTCA 490
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        209 AGGGGCATTACTATGAATGGCTTCCACGATATAATGATCCTGCTGCCCCGTGTGCACTCA 268

Qy        491 AGTGTCATGCACAAGGACAAAACTTGGTGGTGGAGCTGGCACCTAAGGTACTGGATGGAA 550
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        269 AGTGTCATGCACAAGGACAAAACTTGGTGGTGGAGCTGGCACCTAAGGTACTGGATGGAA 328

Qy        551 CTCGTTGCAACACGGACTCCTTGGACATGTGTATCAGTGGCATCTGTCAGGCAGTGGGCT 610
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        329 CTCGTTGCAACACGGACTCCTTGGACATGTGTATCAGTGGCATCTGTCAGGCAGTGGGCT 388

Qy        611 GCGATCGGCAACTGGGAAGCAATGCCAAGGAGGACAACTGTGGAGTCTGTGCCGGCGATG 670
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        389 GCGATCGGCAACTGGGAAGCAATGCCAAGGAGGACAACTGTGGAGTCTGTGCCGGCGATG 448

Qy        671 GCTCCACCTGCAGGCTTGTACGGGGACAATCAAAGTCACACGTTTCTCCTGAAAAAAGAG 730
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        449 GCTCCACCTGCAGGCTTGTACGGGGACAATCAAAGTCACACGTTTCTCCTGAAAAAAGAG 508

Qy        731 AAGAAAATGTAATTGCTGTTCCTTTGGGAAGTCGAAGTGTGAGAATTACAGTGAAAGGAC 790
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        509 AAGAAAATGTAATTGCTGTTCCTTTGGGAAGTCGAAGTGTGAGAATTACAGTGAAAGGAC 568

Qy        791 CTGCCCACCTCTTTATTGAATCAAAAACACTTCAAGGAAGCAAAGGAGAACACAGCTTTA 850
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        569 CTGCCCACCTCTTTATTGAATCAAAAACACTTCAAGGAAGCAAAGGAGAACACAGCTTTA 628

Qy        851 ACAGCCCCGGCGTCTTTGTCGTAGAAAACACAACAGTGGAATTTCAGAGGGGCTCCGAGA 910
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        629 ACAGCCCCGGCGTCTTTGTCGTAGAAAACACAACAGTGGAATTTCAGAGGGGCTCCGAGA 688

Qy        911 GGCAAACTTTTAAGATTCCAGGACCTCTGATGGCTGATTTCATCTTCAAGACCAGGTACA 970
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        689 GGCAAACTTTTAAGATTCCAGGACCTCTGATGGCTGATTTCATCTTCAAGACCAGGTACA 748

Qy        971 CTGCAGCCAAAGACAGCGTGGTTCAGTTCTTCTTTTACCAGCCCATCAGTCATCAGTGGA 1030
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        749 CTGCAGCCAAAGACAGCGTGGTTCAGTTCTTCTTTTACCAGCCCATCAGTCATCAGTGGA 808

Qy       1031 GACAAACTGACTTCTTTCCCTGCACTGTGACGTGTGGAGGAGGTTATCAGCTCAATTCTG 1090
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        809 GACAAACTGACTTCTTTCCCTGCACTGTGACGTGTGGAGGAGGTTATCAGCTCAATTCTG 868

Qy       1091 CTGAATGTGTGGATATCCGCTTGAAGAGGGTAGTTCCTGACCATTATTGTCACTACTACC 1150
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        869 CTGAATGTGTGGATATCCGCTTGAAGAGGGTAGTTCCTGACCATTATTGTCACTACTACC 928

Qy       1151 CTGAAAATGTAAAACCAAAACCAAAACTGAAGGAATGCAGCATGGATCCCTGCCCATCAA 1210
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        929 CTGAAAATGTAAAACCAAAACCAAAACTGAAGGAATGCAGCATGGATCCCTGCCCATCAA 988

Qy       1211 GTGATGGATTTAAAGAGATAATGCCCTATGACCACTTCCAACCTCTTCCTCGCTGGGAAC 1270
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        989 GTGATGGATTTAAAGAGATAATGCCCTATGACCACTTCCAACCTCTTCCTCGCTGGGAAC 1048

Qy       1271 ATAATCCTTGGACTGCATGTTCCGTGTCCTGTGGAGGAGGGATTCAGAGACGGAGCTTTG 1330
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1049 ATAATCCTTGGACTGCATGTTCCGTGTCCTGTGGAGGAGGGATTCAGAGACGGAGCTTTG 1108

Qy       1331 TGTGTGTAGAGGAATCCATGCATGGAGAGATATTGCAGGTGGAAGAATGGAAGTGCATGT 1390
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1109 TGTGTGTAGAGGAATCCATGCATGGAGAGATATTGCAGGTGGAAGAATGGAAGTGCATGT 1168

Qy       1391 ACGCACCCAAACCCAAGGTTATGCAAACTTGTAATCTGTTTGATTGCCCCAAGTGGATTG 1450
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1169 ACGCACCCAAACCCAAGGTTATGCAAACTTGTAATCTGTTTGATTGCCCCAAGTGGATTG 1228

Qy       1451 CCATGGAGTGGTCTCAGTGCACAGTGACTTGTGGCCGAGGGTTACGGTACCGGGTTGTTC 1510
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1229 CCATGGAGTGGTCTCAGTGCACAGTGACTTGTGGCCGAGGGTTACGGTACCGGGTTGTTC 1288

Qy       1511 TGTGTATTAACCACCGCGGAGAGCATGTTGGGGGCTGCAATCCACAACTGAAGTTACACA 1570
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1289 TGTGTATTAACCACCGCGGAGAGCATGTTGGGGGCTGCAATCCACAACTGAAGTTACACA 1348

Qy       1571 TCAAAGAAGAATGTGTCATTCCCATCCCGTGTTATAAACCAAAAGAAAAAAGTCCAGTGG 1630
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1349 TCAAAGAAGAATGTGTCATTCCCATCCCGTGTTATAAACCAAAAGAAAAAAGTCCAGTGG 1408

Qy       1631 AAGCAAAATTGCCTTGGCTGAAACAAGCACAAGAACTAGAAGAGACCAGAATAGCAACAG 1690
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1409 AAGCAAAATTGCCTTGGCTGAAACAAGCACAAGAACTAGAAGAGACCAGAATAGCAACAG 1468

Qy       1691 AAGAACCAACGTTCATTCCAGAACCCTGGTCAGCCTGCAGTACCACGTGTGGCCCAGGTG 1750
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1469 AAGAACCAACGTTCATTCCAGAACCCTGGTCAGCCTGCAGTACCACGTGTGGCCCAGGTG 1528

Qy       1751 TGCAGGTCCGCGAGGTGAAGTGCCGTGTGCTCCTCACATTCACGCAGACTGAGACTGAGC 1810
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1529 TGCAGGTCCGCGAGGTGAAGTGCCGTGTGCTCCTCACATTCACGCAGACTGAGACTGAGC 1588

Qy       1811 TGCCCGAGGAAGAGTGTGAAGGCCCCAAGCTGCCCACCGAACGGCCCTGCCTCCTGGAAG 1870
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1589 TGCCCGAGGAAGAGTGTGAAGGCCCCAAGCTGCCCACCGAACGGCCCTGCCTCCTGGAAG 1648

Qy       1871 CATGTGATGAGAGCCCGGCCTCCCGAGAGCTAGACATCCCTCTCCCTGAGGACAGTGAGA 1930
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1649 CATGTGATGAGAGCCCGGCCTCCCGAGAGCTAGACATCCCTCTCCCTGAGGACAGTGAGA 1708

Qy       1931 CGACTTACGACTGGGAGTACGCTGGGTTCACCCCTTGCACAGCAACATGCTTGGGAGGCC 1990
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1709 CGACTTACGACTGGGAGTACGCTGGGTTCACCCCTTGCACAGCAACATGCTTGGGAGGCC 1768

Qy       1991 ATCAAGAAGCCATAGCAGTGTGCTTACATATCCAGACCCAGCAGACAGTCAATGACAGCT 2050
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1769 ATCAAGAAGCCATAGCAGTGTGCTTACATATCCAGACCCAGCAGACAGTCAATGACAGCT 1828

Qy       2051 TGTGTGATATGGTCCACCGTCCTCCAGCCATGAGCCAGCCCTGTAACACAGAGCCCTGTC 2110
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1829 TGTGTGATATGGTCCACCGTCCTCCAGCCATGAGCCAGCCCTGTAACACAGAGCCCTGTC 1888

Qy       2111 CCCCCAGGTGGCATGTGGGCTCTTGGGGGCCCTGCTCAGCTACCTGTGGAGTTGGAATTC 2170
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1889 CCCCCAGGTGGCATGTGGGCTCTTGGGGGCCCTGCTCAGCTACCTGTGGAGTTGGAATTC 1948

Qy       2171 AGACCCGAGATGTGTACTGCCTGCACCCAGGGGAGACCCCTGCCCCTCCTGAGGAGTGCC 2230
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1949 AGACCCGAGATGTGTACTGCCTGCACCCAGGGGAGACCCCTGCCCCTCCTGAGGAGTGCC 2008

Qy       2231 GAGATGAAAAGCCCCATGCTTTACAAGCATGCAATCAGTTTGACTGCCCTCCTGGCTGGC 2290
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2009 GAGATGAAAAGCCCCATGCTTTACAAGCATGCAATCAGTTTGACTGCCCTCCTGGCTGGC 2068

Qy       2291 ACATTGAAGAATGGCAGCAGTGTTCCAGGACTTGTGGCGGGGGAACTCAGAACAGAAGAG 2350
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2069 ACATTGAAGAATGGCAGCAGTGTTCCAGGACTTGTGGCGGGGGAACTCAGAACAGAAGAG 2128

Qy       2351 TCACCTGTCGGGCAGCTGCTAACGGATGGCAGCTTTTGAATCTCTCAGATGAATTGTGCC 2410
```

```
              [illegible]
Db       2129 [illegible] 2188

Qy       2411 [illegible] 2470
              [illegible]
Db       2189 [illegible] 2248

Qy       2471 [illegible] 2490
              [illegible]
Db       2249 [illegible] 2268


RESULT 11
AAD14967
ID   AAD14967 standard; cDNA; 2175 BP.
XX
AC   AAD14967;
XX
DT   21-NOV-2001  (first entry)
XX
DE   Novel human protein (NHP) cDNA #4.
XX
KW   Novel human protein; NHP; thrombospondin; gene therapy; cancer;
KW   cytostatic; antisense therapy; angiogenesis; biological disorder; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..2175
FT                   /*tag=  a
FT                   /product= "Novel human protein"
XX
PN   WO200161011-A2.
XX
PD   23-AUG-2001.
XX
PF   15-FEB-2001; 2001WO-US05290.
XX
PR   17-FEB-2000; 2000US-0183282.
XX
PA   (LEXI-) LEXICON GENETICS INC.
XX
PI   Donoho G,  Scoville J,  Turner CA,  Friedrich G,  Zambrowicz B;
PI   Sands AT;
XX
DR   WPI; 2001-514776/56.
DR   P-PSDB; AAE07866.
XX
PT   An isolated nucleic acid encoding a thrombospondin useful as a
PT   hybridization probe and gene therapy treatments of cancer -
XX
PS   Disclosure; Page 38; 59pp; English.
XX
CC   The invention relates to novel human polynucleotides encoding proteins
CC   that share sequence similarity with animal proteins having thrombospondin
CC   repeats. NHPs are expressed in human cell lines, pituitary, lymph nodes,
CC   prostate, testis, adrenal gland, uterus, foetal kidney, foetal lung and
CC   gene trapped human cells. NHPs are thrombospondins useful for treating
CC   biological disorders involving angiogenesis, cancer and development and
CC   also in pharmacogenomic applications. NHPs are useful as a hybridisation
CC   probe for screening libraries, assessing gene expression patterns and
CC   also in gene therapy. Proteins having thrombospondin repeats act as
CC   receptors, secreted extracellular matrix proteins and proteases. The
CC   present sequence is a cDNA encoding novel human protein (NHP).
XX
SQ   Sequence 2175 BP; 583 A; 532 C; 575 G; 485 T; 0 other;

  Query Match             41.7%;  Score [illegible];  DB 22;  Length 2175;
  Best Local Similarity   [illegible];  Pred. No. 0;
  Matches [illegible];  Conservative    0;  Mismatches   16;  Indels   [illegible];  Gaps   [illegible];

Qy          1 [illegible] 60
              [illegible]
Db          1 [illegible] 60

Qy         61 [illegible] 120
              [illegible]
Db         61 [illegible] 120

Qy        121 [illegible] 180
              [illegible]
Db        121 [illegible] 180

Qy        181 [illegible] 240
              [illegible]
Db        181 [illegible] 240

Qy        241 [illegible] 300
              [illegible]
Db        241 [illegible] 300

Qy        301 [illegible] 360
              [illegible]
Db        301 [illegible] 360

Qy        361 [illegible] 420
              [illegible]
Db        361 [illegible] 420

Qy        421 [illegible] 480
              [illegible]
Db        421 [illegible] 480

Qy        481 [illegible] 540
              [illegible]
Db        481 [illegible] 540

Qy        541 [illegible] 600
              [illegible]
Db        541 [illegible] 600

Qy        601 [illegible] 660
              [illegible]
Db        601 [illegible] 660

Qy        661 [illegible] 720
              [illegible]
Db        661 [illegible] 720

Qy        721 GAAAAAAGAGAAGAAAATGTAATTGCTGTTCCTTTGGGAAGTCGAAGTGTGAGAATTACA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GAAAAAAGAGAAGAAAATGTAATTGCTGTTCCTTTGGGAAGTCGAAGTGTGAGAATTACA 780

Qy        781 [illegible] 840
              [illegible]
Db        781 [illegible] 840

Qy        841 [illegible] 900
              [illegible]
Db        841 [illegible] 900

Qy        901 [illegible] 960
              [illegible]
Db        901 [illegible] 960

Qy        961 [illegible] 1020
              [illegible]
Db        961 [illegible] 1020

Qy       1021 [illegible] 1080
              [illegible]
Db       1021 [illegible] 1080

Qy       1081 [illegible] 1140
              [illegible]
Db       1081 [illegible] 1140
```

```
Qy     1141 CACTACTACCCTGAAAATGTAAAACCAAAACCAAAACTGAAGGAATGCAGCATGGATCCC 1200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1141 CACTACTACCCTGAAAATGTAAAACCAAAACCAAAACTGAAGGAATGCAGCATGGATCCC 1200

Qy     1201 TGCCCATCAAGTGATGGATTTAAAGAGATAATGCCCTATGACCACTTCCAACCTCTTCCT 1260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1201 TGCCCATCAAGTGATGGATTTAAACAGATAATGCCCTATGACCACTTCCAACCTCTTCCT 1260

Qy     1261 CGCTGGGAACATAATCCTTGGACTGCATGTTCCGTGTCCTGTGGAGGAGGGATTCAGAGA 1320
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1261 CGCTGGGAACATAATCCTTGGACTGCATGTTCCGTGTCCTGTGGAGGAGGGATTCAGAGA 1320

Qy     1321 CGGAGCTTTGTGTGTGTAGAGGAATCCATGCATGGAGAGATATTGCAGGTGGAAGAATGG 1380
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1321 CGGAGCTTTGTGTGTGTAGAGGAATCCATGCATGGAGAGATATTGCAGGTGGAAGAATGC 1380

Qy     1381 AAGTGCATGTACGCACCCAAACCCAAGGTTATGCAAACTTGTAATCTGTTTGATTGCCCC 1440
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1381 AAGTGCATGTACGCACCCAAACCCAAGGTTATGCAAACTTGTAATCTGTTTGATTGCCCC 1440

Qy     1441 AAGTGGATTGCCATGGAGTGGTCTCAGTGCACAGTGACTTGTGGCCGAGGGTTACGGTAC 1500
            ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db     1441 AAGTGGATTGCCATGGAGTGGTCTCAGTGCACAGTGACTTGTGGCCGAGGGTTACGGTAC 1500

Qy     1501 CGGGTTGTTCTGTGTATTAACCACCGCGGAGAGCATGTTGGGGGCTGCAATCCACAACTG 1560
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1501 CGGGTTGTTCTGTGTATTAACCACCGCGGAGAGCATGTTGGCGGCTGCAATCCACAACTG 1560

Qy     1561 AAGTTACACATCAAAGAAGAATGTGTCATTCCCATCCCGTGTTATAAACCAAAAGAAAAA 1620
            ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db     1561 AAGTTACACATCAAAGAAGAATGTGTCATTCCCATCCCGTGTTATAAACCAAAAGAAAAA 1620

Qy     1621 AGTCCAGTGGAAGCAAAATTGCCTTGGCTGAAACAAGCACAAGAACTAGAAGAGACCAGA 1680
            ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db     1621 AGTCCAGTGGAAGCAAAATTGCCTTGGCTGAAACAAGCACAAGAACTAGAAGAGACCAGA 1680

Qy     1681 ATAGCAACAGAAGAACCAACGTTCATTCCAGAACCCTGGTCAGCCTGCAGTACCACGTGT 1740
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1681 ATAGCAACAGAAGAACCAACGTTCATTCCAGAACCCTGGTCAGCCTGCAGTACCACGTGT 1740

Qy     1741 GGGCCAGGTGTGCAGGTCCGCGAGGTGAAGTGCCGTGTGCTCCTCACATTCACGCAGACT 1800
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1741 GGGCCAGGTGTGCAGGTCCGCGAGGTGAAGTGCCGTGTGCTCCTCACATTCACGCAGACT 1800

Qy     1801 GAGACTGAGCTGCCCGAGGAAGAGTGTGAAGGCCCCAAGCTGCCCACCGAACGGCCCTGC 1860
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1801 GAGACTGAGCTGCCCGAGGAAGAGTGTGAAGGCCCCAAGCTGCCCACCGAACGGCCCTGC 1860

Qy     1861 CTCCTGGAAGCATGTGATGAGAGCCCGGCCTCCCGAGAGCTAGACATCCCTCTCCCTGAG 1920
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1861 CTCCTGGAAGCATGTGATGAGAGCCCGGCCTCCCGAGAGCTAGACATCCCTCTCCCTGAG 1920

Qy     1921 GACAGTGAGACGACTTACGACTGGGAGTACGCTGGCTTCACCCCTTGCACAGCAACATGC 1980
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1921 GACAGTGAGACGACTTACGACTGGGAGTACGCTGGCTTCACCCCTTGCACAGCAACATGC 1980

Qy     1981 TTGGGAGGCCATCAAGAAGCCATAGCAGTGTGCTTACATATCCAGACCCAGCAGACAGTC 2040
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1981 TTGGGAGGCCATCAAGAAGCCATAGCAGTGTGCTTACATATCCAGACCCAGCAGACAGTC 2040

Qy     2041 AATGACAGCTTGTGTGATATGGTCCACCGTCCTCCAGCCATGAGCCAGGCCTGTAACACA 2100
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2041 AATGACAGCTTGTGTGATATGGTCCACCGTCCTCCAGCCATGAGCCAGGCCTGTAACACA 2100

Qy     2101 GAGCCCTGTCCCCCCAGGTGGCATGTGGGCTCTTG 2135
            |||||||||||||||||| |  |    |   ||||
Db     2101 GAGCCCTGTCCCCCCAGGAGAGAGCCAGCAGCTTG 2135


RESULT 12
AAD14368
ID   AAD14368 standard; cDNA; 1953 BP.
XX
AC   AAD14368;
XX
DT   01 NOV 2001  (first entry)
XX
DE   Novel human protein (NHP) cDNA #5.
XX
KW   Novel human protein, NHP, thrombospondin, gene therapy; cancer;
KW   cytostatic; antisense therapy; angiogenesis; biological disorder; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1953
FT                   /*tag   a
FT                   /product  "Novel human protein"
XX
PN   WO200161011-A2.
XX
PD   23-AUG-2001.
XX
PF   15-FEB-2001; 2001WO-US05290.
XX
PR   17-FEB-2000; 2000US-0183282.
XX
PA   (LEXI ) LEXICON GENETICS INC.
XX
PI   Donoho G,  Scoville J,  Turner CA,  Friedrich G,  Zambrowicz B;
PI   Sands AT;
XX
DR   WPI; 2001-514776/56.
DR   P-PSDB; AAE07867.
XX
PT   An isolated nucleic acid encoding a thrombospondin useful as a
PT   hybridization probe and gene therapy treatments of cancer -
XX
PS   Disclosure; Page 40-41; 56pp; English.
XX
CC   The invention relates to novel human polynucleotides encoding proteins
CC   that share sequence similarity with animal proteins having thrombospondin
CC   repeats. NHPs are expressed in human cell lines, pituitary, lymph nodes,
CC   prostrate, testis, adrenal gland, uterus, foetal kidney, foetal lung and
CC   gene trapped human cells. NHPs are thrombospondins useful for treating
CC   biological disorders involving angiogenesis, cancer and development and
CC   also in pharmacogenomic applications. NHPs are useful as a hybridisation
CC   probe for screening libraries, assessing gene expression patterns and
CC   also in gene therapy. Proteins having thrombospondin repeats act as
CC   receptors, secreted extracellular matrix proteins and proteases. The
CC   present sequence is a cDNA encoding novel human protein (NHP).
XX
SQ   Sequence 1953 BP; 530 A; 480 C; 513 G; 430 T; 0 other;

  Query Match             35.6%;  Score 1809;  DB 22;  Length 1953;
  Best Local Similarity   99.5%;  Pred. No. 0;
  Matches 1815;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy      311 CTGGAAGGAATTGTGAAGGGCAGAACATTCGGTACAAGACATGCAGCAATCATGACTGCC 370
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       89 CTGGAAGGAATTGTGAAGGGCAGAACATTCGGTACAAGACATGCAGCAATCATGACTGCC 148

Qy      371 CTCCAGATGCAGAAGATTTCAGAGCCCAGCAGTGCTCAGCCTACAATGATGTCCAGTATC 430
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      149 CTCCAGATGCAGAAGATTTCAGAGCCCAGCAGTGCTCAGCCTACAATGATGTCCAGTATC 208

Qy      431 AGGGGCATTACTATGAATGGCTTCCACGATATAATGATCCTGCTGCCCCGTGTGCACTCA 490
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      209 AGGGGCATTACTATGAATGGCTTCCACGATATAATGATCCTGCTGCCCCGTGTGCACTCA 268

Qy      491 AGTGTCATGCACAAGGACAAAACTTGGTGGTGGAGCTGGCACCTAAGGTACTGGATGGAA 550
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      269 AGTGTCATGCACAAGGACAAAACTTGGTGGTGGAGCTGGCACCTAAGGTACTGGATGGAA 328

Qy      551 CTCGTTGCAACACGGACTCCTTGGACATGTGTATCAGTGGCATCTGTCAGGCAGTGGGCT 610
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        329 CTCGTTGCAACACGGACTCTTGGACATGTGTATCAGTGGCATCTGTCAGGCAGTGGGCT 388

Qy        611 GCGATCGGCAACTGGGAAGCAATGCCAAGGAGGACAACTGTGGAGTCTGTGCCGGGGATG 670
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        389 GCGATCGGCAACTGGGAAGCAATGCCAAGGAGGACAACTGTGGAGTCTGTGCCGGGGATG 448

Qy        671 GCTCCACCTGCAGCCTTGTACGGGGACAATCAAAGTCACAGTTTCTTCTGAAAAAAGAG 730
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        449 GCTCCACCTGCAGCCTTGTACGGGGACAATCAAAGTCACAGTTTCTTCTGAAAAAAGAG 508

Qy        731 AAGAAAATGTAATTGCTGTTCCTTTGGGAAGTCGAAGTGTGAGAATTACAGTGAAAGGAC 790
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        509 AAGAAAATGTAATTGCTGTTCCTTTGGGAAGTCGAAGTGTGAGAATTACAGTGAAAGGAC 568

Qy        791 CTGCCCACCTCTTTATTGAATCAAAAACACTTCAAGGAAGCAAAGGAGAACACAGCTTTA 850
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        569 CTGCCCACCTCTTTATTGAATCAAAAACACTTCAAGGAAGCAAAGGAGAACACAGCTTTA 628

Qy        851 ACAGCCCCGGCGTCTTTGTCGTAGAAAACACAACAGTGGAATTTCAGAGGGGTCCGAGA 910
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        629 ACAGCCCCGGCGTCTTTGTCGTAGAAAACACAACAGTGGAATTTCAGAGGGGTCCGAGA 688

Qy        911 CCCAAACTTTTAAGATTCTAGGACTCTGATGGCTGATTCATCTGAAGACCAGGTATA 970
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        689 CCCAAACTTTTAAGATTCTAGGACTCTGATGGCTGATTCATCTGAAGACCAGGTATA 748

Qy        971 CTGCAGCCAAAGACAGCCCCCTTCAGTTCTTCTTTACCAGCCATCAGCATCAGCTGGA 1030
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        749 CTGCAGCCAAAGACAGCCCCCTTCAGTTCTTCTTTACCAGCCATCAGCATCAGCTGGA 808

Qy       1031 GACAAACTGACTTCTTTCCTTGGCACTGGACGTGTGGAGGAGGTTATCAGCTCAATTCTG 1090
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        809 GACAAACTGACTTCTTTCCTTGGCACTGGACGTGTGGAGGAGGTTATCAGCTCAATTCTG 868

Qy       1091 CTGAATGTGTGGATATCCGCTTGAAGAGGGTAGTTCCTGACCATTATTGTCACTACTACC 1150
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        869 CTGAATGTGTGGATATCCGCTTGAAGAGGGTAGTTCCTGACCATTATTGTCACTACTACC 928

Qy       1151 CTGAAAATGTAAAACCAAAACCAAAACTGAAGGAATGCAGCATGGATCCCTGCCCATCAA 1210
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        929 CTGAAAATGTAAAACCAAAACCAAAACTGAAGGAATGCAGCATGGATCCCTGCCCATCAA 988

Qy       1211 GTGATGGATTTAAAGAGATAATGCCCTATGACCACTTCCAACCTCTTCCTCGCTGGGAAC 1270
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        989 GTGATGGATTTAAAGAGATAATGCCCTATGACCACTTCCAACCTCTTCCTCGCTGGGAAC 1048

Qy       1271 ATAATCCTTGGACTGCATGCTCCGTTTCCTGTGGAGGAGGGATTCAGAGACGGAGCTTTG 1330
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1049 ATAATCCTTGGACTGCATGCTCCGTTTCCTGTGGAGGAGGGATTCAGAGACGGAGCTTTG 1108

Qy       1331 TGTGTGTAGAGGAATCCATGCATGGAGAGATATTGCAGGTGGAAGAATGGAAGTGCATGT 1390
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1109 TGTGTGTAGAGGAATCCATGCATGGAGAGATATTGCAGGTGGAAGAATGGAAGTGCATGT 1168

Qy       1391 ACGCACCCAAACCCAAGGTTATGCAAACTTGTAATCTGTTTGATTGCCCCAAGTGGATTG 1450
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1169 ACGCACCCAAACCCAAGGTTATGCAAACTTGTAATCTGTTTGATTGCCCCAAGTGGATTG 1228

Qy       1451 CCATGGAGTGGTCTCAGTGCACAGTGACTTGTGGCCGAGGGTTACGGTACCGGGTTGTTC 1510
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1229 CCATGGAGTGGTCTCAGTGCACAGTGACTTGTGGCCGAGGGTTACGGTACCGGGTTGTTC 1288

Qy       1511 TGTGTATTAACCACCGCGGAGAGCATGTTGGGGGCTGCAATCCACAACTGAAGTTACACA 1570
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1289 TGTGTATTAACCACCGCGGAGAGCATGTTGGGGGCTGCAATCCACAACTGAAGTTACACA 1348

Qy       1571 TCAAAGAAGAATGTGTCATTCCCATCCCGTGTTATAAACCAAAAGAAAAAAGTCCAGTGG 1630
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1349 TCAAAGAAGAATGTGTCATTCCCATCCCGTGTTATAAACCAAAAGAAAAAAGTCCAGTGG 1408

Qy       1631 AAGCAAAAATTGCCTGGCTGAAACAAGCACAAGAACTAGAAGAGACCAGAATAGCAACAG 1690
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1409 AAGCAAAAATTGCCTGGCTGAAACAAGCACAAGAACTAGAAGAGACCAGAATAGCAACAG 1468

Qy       1691 AAGAACCAACGTTCATTCCAGAACCCTGGTCAGCCTGCAGTACCACGTGTGGGCCAGGTG 1750
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1469 AAGAACCAACGTTCATTCCAGAACCCTGGTCAGCCTGCAGTACCACGTGTGGGCCAGGTG 1528

Qy       1751 TGCAGGTCCGCGAGGTGAAGTGCCGTGTGCTCCTCACATTCACGCAGACTGAGACTGAGC 1810
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1529 TGCAGGTCCGCGAGGTGAAGTGCCGTGTGCTCCTCACATTCACGCAGACTGAGACTGAGC 1588

Qy       1811 TGCCCGAGGAAGAGTGTGAAGGCCCCAAGCTGCCCACCGAACGGCCCTGCCTCCTGGAAG 1870
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1589 TGCCCGAGGAAGAGTGTGAAGGCCCCAAGCTGCCCACCGAACGGCCCTGCCTCCTGGAAG 1648

Qy       1871 CATGTGATGAGAGCCCGGCCTCCCGAGAGCTAGACATCCCTCTCCCTGAGGACAGTGAGA 1930
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1649 CATGTGATGAGAGCCCGGCCTCCCGAGAGCTAGACATCCCTCTCCCTGAGGACAGTGAGA 1708

Qy       1931 CGACTTACGACTGGGAGTACGCTGGGTTCACCCCTTGCACAGCAACATGCTTGGGAGGCC 1990
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1709 CGACTTACGACTGGGAGTACGCTGGGTTCACCCCTTGCACAGCAACATGCTTGGGAGGCC 1768

Qy       1991 ATCAAGAAGCCATAGCAGTGTGCTTACATATCCAGACCCAGCAGACAGTCAATGACAGCT 2050
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1769 ATCAAGAAGCCATAGCAGTGTGCTTACATATCCAGACCCAGCAGACAGTCAATGACAGCT 1828

Qy       2051 TGTGTGATATGGTCCACCGTCCTCCAGCCATGAGCCAGGCCTGCAATACAGAGCCCTGTC 2110
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1829 TGTGTGATATGGTCCACCGTCCTCCAGCCATGAGCCAGGCCTGCAATACAGAGCCCTGTC 1888

Qy       2111 CCCCCAGGTGGCATGTGGGCTCTTG 2135
              ||||||||  |  |     |    ||||
Db       1889 CCCCCAGGAGAGAGCCAGCAGCTTG 1913


RESULT 13
AAD14365
ID   AAD14365 standard; cDNA; 1341 BP.
XX
AC   AAD14365;
XX
DT   01-NOV-2001  (first entry)
XX
DE   Novel human protein (NHP) cDNA #2.
XX
KW   Novel human protein; NHP; thrombospondin; gene therapy; cancer;
KW   cytostatic; antisense therapy; angiogenesis; biological disorder; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1341
FT                   /*tag=  a
FT                   /product= "Novel human protein"
XX
PN   WO200161011-A2.
XX
PD   23-AUG-2001.
XX
PF   15-FEB-2001; 2001WO-US-04790.
XX
PR   17-FEB-2000; 2000US-0183282.
XX
PA   (LEXI-) LEXICON GENETICS INC.
XX
PI   Donoho G,  Scoville J,  Turner CA,  Friedrich G,  Zambrowicz B;
PI   Sands AT;
XX
DR   WPI; 2001-514279/56.
DR   P-PSDB; AAE07864.
XX
PT   An isolated nucleic acid encoding a thrombospondin useful as a
PT   hybridization probe and gene therapy treatments of cancer
```

```
XX
PS   Disclosure; Page 35; 56pp; English.
XX
CC   The invention relates to novel human polynucleotides encoding proteins
CC   that share sequence similarity with animal proteins having thrombospondin
CC   repeats. NHPs are expressed in human cell lines, pituitary, lymph nodes,
CC   prostrate, testis, adrenal gland, uterus, foetal kidney, foetal lung and
CC   gene trapped human cells. NHPs are thrombospondins useful for treating
CC   biological disorders involving angiogenesis, cancer and development and
CC   also in pharmacogenomic applications. NHPs are useful as a hybridisation
CC   probe for screening libraries, assessing gene expression patterns and
CC   also in gene therapy. Proteins having thrombospondin repeats act as
CC   receptors, secreted extracellular matrix proteins and proteases. The
CC   present sequence is a cDNA encoding novel human protein (NHP).
XX
SQ   Sequence 1341 BP; 357 A; 315 C; 349 G; 320 T; 0 other;

  Query Match             26.1%;  Score 1327;  DB 22;  Length 1341;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1339;  Conservative    0;  Mismatches    0;  Indels    2;  Gaps    1;

Qy          1 ATGGCTTCCTGGACGAGCCCCTGGTGGGTGCTGATAGGGATGGTCTTCATGCACTCTCCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGCTTCCTGGACGAGCCCCTGGTGGGTGCTGATAGGGATGGTCTTCATGCACTCTCCC 60

Qy         61 CTCCCGCAGACCACAGCTGAGAAATCTCCTGGAGCCTATTTCCTTCCCGAGTTTGCACTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTCCCGCAGACCACAGCTGAGAAATCTCCTGGAGCCTATTTCCTTCCCGAGTTTGCACTT 120

Qy        121 TCTCCTCAGGGAAGTTTTCTGGAAGACACAACAGGGGAGCAGTTCCTCACTTATCGCTAT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCTCCTCAGGGAAGTTTTCTGGAAGACACAACAGGGGAGCAGTTCCTCACTTATCGCTAT 180

Qy        181 GATGACCAGACCTCAAGAAACACTCGTTCAGATGAAGACAAAGATGGCAACTGGGATGCT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GATGACCAGACCTCAAGAAACACTCGTTCAGATGAAGACAAAGATGGCAACTGGGATGCT 240

Qy        241 TGGGGCGACTGGAGTGACTGCTCCCGGACCTGTGGGGGAGGAGCATCATATTCTCTGCGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGGGGCGACTGGAGTGACTGCTCCCGGACCTGTGGGGGAGGAGCATCATATTCTCTGCGG 300

Qy        301 AGATGTTTGACTGGAAGGAATTGTGAAGGGCAGAACATTCGGTACAAGACATGCAGCAAT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGATGTTTGACTGGAAGGAATTGTGAAGGGCAGAACATTCGGTACAAGACATGCAGCAAT 360

Qy        361 CATGACTGCCCTCCAGATGCAGAAGATTTCAGAGCCCAGCAGTGCTCAGCCTACAATGAT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CATGACTGCCCTCCAGATGCAGAAGATTTCAGAGCCCAGCAGTGCTCAGCCTACAATGAT 420

Qy        421 GTCCAGTATCAGGGGCATTACTATGAATGGCTTCCACGATATAATGATCCTGCTGCCCCG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GTCCAGTATCAGGGGCATTACTATGAATGGCTTCCACGATATAATGATCCTGCTGCCCCG 480

Qy        481 TGTGCACTCAAGTGTCATGCACAAGGACAAAACTTCGTGGTGGAGCTGGCACCTAAGGTA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TGTGCACTCAAGTGTCATGCACAAGGACAAAACTTCGTGGTGGAGCTGGCACCTAAGGTA 540

Qy        541 CTGGATGGAACTCGTTGCAACACGGACTCCTTGGACATGTGTATCAGTGGCATCTGTCAG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTGGATGGAACTCGTTGCAACACGGACTCCTTGGACATGTGTATCAGTGGCATCTGTCAG 600

Qy        601 GCAGTGGGCTGCGATCGGCAACTGGGAAGCAATGCCAAGGAGGACAACTGTGGAGTCTGT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GCAGTGGGCTGCGATCGGCAACTGGGAAGCAATGCCAAGGAGGACAACTGTGGAGTCTGT 660

Qy        661 GCCGGCGATGGCTCCACCTGCAGGCTTGTACGGGGACAATCAAAGTCACACGTTTCTCCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GCCGGCGATGGCTCCACCTGCAGGCTTGTACGGGGACAATCAAAGTCACACGTTTCTCCT 720

Qy        721 GAAAAAAGAGAAGAAAATGTAATTGCTGTTCCTTTGGGAAGTCGAAGTGTGAGAATTACA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GAAAAAAGAGAAGAAAATGTAATTGCTGTTCCTTTGGGAAGTCGAAGTGTGAGAATTACA 780

Qy        781 GTGAAAGGACCTGCCCACCTCTTTATTGAATCAAAAACACTTCAAGGAAGCAAAGGAGAA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GTGAAAGGACCTGCCCACCTCTTTATTGAATCAAAAACACTTCAAGGAAGCAAAGGAGAA 840

Qy        841 CACAGCTTTAACAGCCCCGGCGTCTTTGTCGTAGAAAACACAACAGTGGAATTTCAGAGG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CACAGCTTTAACAGCCCCGGCGTCTTTGTCGTAGAAAACACAACAGTGGAATTTCAGAGG 900

Qy        901 GGCTCCGAGAGGCAAACTTTTAAGATTCCAGGACCTCTGATGGCTGATTTCATCTTCAAG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GGCTCCGAGAGGCAAACTTTTAAGATTCCAGGACCTCTGATGGCTGATTTCATCTTCAAG 960

Qy        961 ACCAGGTACACTGCAGCCAAAGACAGCGTGGTTCAGTTCTTCTTTACCAGCCCATCAGT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 ACCAGGTACACTGCAGCCAAAGACAGCGTGGTTCAGTTCTTCTTTACCAGCCCATCAGT 1020

Qy       1021 CATCAGTGGAGACAAACTGACTTCTTTCCCTGCACTGTGACGTGTGGAGGAGGTTATCAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CATCAGTGGAGACAAACTGACTTCTTTCCCTGCACTGTGACGTGTGGAGGAGGTTATCAG 1080

Qy       1081 CTCAATTCTGCTGAATGTGTGGATATCCGCTTGAAGAGGGTAGTTCCTGACCATTATTGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CTCAATTCTGCTGAATGTGTGGATATCCGCTTGAAGAGGGTAGTTCCTGACCATTATTGT 1140

Qy       1141 CACTACTACCCTGAAAATGTAAAACCAAAACCAAAACTGAAGGAATGCAGCATGGATCCC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CACTACTACCCTGAAAATGTAAAACCAAAACCAAAACTGAAGGAATGCAGCATGGATCCC 1200

Qy       1201 TGCCCATCAAGTGATGGATTTAAAGAGATAATGCCCTATGACCACTTCCAACCTCTTCCT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TGCCCATCAAGTGATGGATTTAAAGAGATAATGCCCTATGACCACTTCCAACCTCTTCCT 1260

Qy       1261 C--GCTGGGAACATAATCCTTGGACTGCATGTTCCGTGTCCTGTGGAGGAGGGATTCAGA 1318
              |  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CGAGCTGGGAACATAATCCTTGGACTGCATGTTCCGTGTCCTGTGGAGGAGGGATTCAGA 1320

Qy       1319 GACGGAGCTTTGTGTGTGTAG 1339
              |||||||||||||||||||||
Db       1321 GACGGAGCTTTGTGTGTGTAG 1341


RESULT 14
AAD14366
ID   AAD14366 standard; cDNA; 1119 BP.
XX
AC   AAD14366;
XX
DT   01 NOV 2001  (first entry)
XX
DE   Novel human protein (NHP) cDNA #3.
XX
KW   Novel human protein, NHP, thrombospondin, gene therapy, cancer;
KW   cytostatic; antisense therapy, angiogenesis, biological disorder; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1119
FT                   /*tag=  a
FT                   /product  "Novel human protein"
XX
PN   WO200161011-A2.
XX
PD   23 AUG 2001.
XX
PF   15-FEB-2001; 2001WO-US05290.
XX
PR   17-FEB-2000; 2000US-0183282.
XX
PA   (LEXI-) LEXICON GENETICS INC.
XX
PI   Donoho G,  Scoville J,  Turner CA,  Friedrich G,  Zambrowicz B;
PI   Sands AT;
```

```
XX
DR   WPI; 2001-514776/56.
DR   P-PSDB; AAE07865.
XX
PT   An isolated nucleic acid encoding a thrombospondin useful as a
PT   hybridization probe and gene therapy treatments of cancer
XX
PS   Disclosure; Page 36; 56pp; English.
XX
CC   The invention relates to novel human polynucleotides encoding proteins
CC   that share sequence similarity with animal proteins having thrombospondin
CC   repeats. NHPs are expressed in human cell lines, pituitary, lymph nodes,
CC   prostrate, testis, adrenal gland, uterus, foetal kidney, foetal lung and
CC   gene trapped human cells. NHPs are thrombospondins useful for treating
CC   biological disorders involving angiogenesis, cancer and development and
CC   also in pharmacogenomic applications. NHPs are useful as a hybridisation
CC   probe for screening libraries, assessing gene expression patterns and
CC   also in gene therapy. Proteins having thrombospondin repeats act as
CC   receptors, secreted extracellular matrix proteins and proteases. The
CC   present sequence is a cDNA encoding novel human protein (NHP).
XX
SQ   Sequence 1119 BP; 301 A; 264 C; 287 G; 267 T; 0 other;

  Query Match             20.4%;  Score 1017;  DB 22;  Length 1119;
  Best Local Similarity   99.8%;  Pred. No. 1.1e-306;
  Matches 1029;  Conservative    0;  Mismatches    0;  Indels    2;  Gaps    1;

Qy        311 CTGGAAGGAATTGTGAAGGGCAGAACATTCGGTACAAGACATGCAGGAATCATGACTGGT 370
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         89 CTGGAAGGAATTGTGAAGGGCAGAACATTCGGTACAAGACATGCAGGAATCATGACTGGT 148

Qy        371 CTCCAGATGCAGAAGATTTTAGAGCCCAGCAGTGCTCAGCCTACAATGATGTTCAGTATC 430
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        149 CTCCAGATGCAGAAGATTTTAGAGCCCAGCAGTGCTCAGCCTACAATGATGTTCAGTATC 208

Qy        431 AGGGGCATTACTATGAATGGCTTCCACGATATAATGATCCTGCTGCCCCGTGTGCACTCA 490
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        209 AGGGGCATTACTATGAATGGCTTCCACGATATAATGATCCTGCTGCCCCGTGTGCACTCA 268

Qy        491 ACTGTCATGCACAAGGACAAAACTGCGTGCTGGAGCTGGCACCTAAGGTACTGGATGGAA 550
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        269 ACTGTCATGCACAAGGACAAAACTGCGTGCTGGAGCTGGCACCTAAGGTACTGGATGGAA 328

Qy        551 CTCCTTGCAACAGGACTCCTTGGACATGTGTATCAGTGGCATCTGTCAGCCAGTGCCCTT 610
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        329 CTCCTTGCAACAGGACTCCTTGGACATGTGTATCAGTGGCATCTGTCAGCCAGTGCCCTT 388

Qy        611 GCTATCGGCAACTGGGAAGCAATGCTAAGGAGGACAACTGTGGAGTCTGTGCCCGGGATG 670
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        389 GCTATCGGCAACTGGGAAGCAATGCTAAGGAGGACAACTGTGGAGTCTGTGCCCGGGATG 448

Qy        671 GCTCCACTGCAGGCTTTTCACTGGGACAATCAAAGTCACAGCTTTTCTCTGAAAAAAGAC 730
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        449 GCTCCACTGCAGGCTTTTCACTGGGACAATCAAAGTCACAGCTTTTCTCTGAAAAAAGAC 508

Qy        731 AAGAAAATGTAATTGCTCTTCCTTCTTTGGAAGTGAAGTGTGAGAATTACATCAAAGGAC 790
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        509 AAGAAAATGTAATTGCTCTTCCTTCTTTGGAAGTGAAGTGTGAGAATTACATCAAAGGAC 568

Qy        791 CTGCCCACCTCTTTATTGAATCAAAAACACTTCAAGGAAGCAAAGGAGAACACAGCTTTA 850
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        569 CTGCCCACCTCTTTATTGAATCAAAAACACTTCAAGGAAGCAAAGGAGAACACAGCTTTA 628

Qy        851 ACAGCCCCGGCGTCTTTGTCGTAGAAAACACAACAGTGGAATTTCAGAGGGGCTCCGAGA 910
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        629 ACAGCCCCGGCGTCTTTGTCGTAGAAAACACAACAGTGGAATTTCAGAGGGGCTCCGAGA 688

Qy        911 GCCAAATTTTTAAGATTCTAGGACCTTTGATGGTGATTTCATCTTCAAGAATCAGGTACA 970
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        689 GCCAAATTTTTAAGATTCTAGGACCTTTGATGGTGATTTCATCTTCAAGAATCAGGTACA 748

Qy        971 CTGCAGCCAAAGACAGCGTGGTTCAGTTCTTCTTTTACCAGCCCATCAGTTACCAGTGGA 1030
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        749 CTGCAGCCAAAGACAGCGTGGTTCAGTTCTTCTTTTACCAGCCCATCAGTTACCAGTGGA 808

Qy       1031 GACAAACTGACTTCTTTCCCTGCACTGTGACGTGTGGAGGAGGTTATCAGCTTAATTCTG 1090
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        809 GACAAACTGACTTCTTTCCCTGCACTGTGACGTGTGGAGGAGGTTATCAGCTTAATTCTG 868

Qy       1091 CTGAATGTGTGGATATCCGCTTGAAGAGGGTAGTTCCTGACCATTATTGTCACTACTACC 1150
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        869 CTGAATGTGTGGATATCCGCTTGAAGAGGGTAGTTCCTGACCATTATTGTCACTACTACC 928

Qy       1151 CTGAAAATGTAAAACCAAAACCAAAACTGAAGGAATGCAGCATGGATCCCTGCCCATCAA 1210
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        929 CTGAAAATGTAAAACCAAAACCAAAACTGAAGGAATGCAGCATGGATCCCTGCCCATCAA 988

Qy       1211 GTGATGGATTTAAAGAGATAATGCCCTATGACCACTTCCAACCTTTTCCTG--CTGGGA 1268
              |||||||||||||||||||||||||||||||||||||||||||||||||||  ||||||
Db        989 GTGATGGATTTAAAGAGATAATGCCCTATGACCACTTCCAACCTTTTCCTGGACTGGGA 1048

Qy       1269 ACATAATCCTTGGACTGCATGTTCCGTGTCCTGTGGAGGAGGGATTCAGAGACGGAGCTT 1328
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1049 ACATAATCCTTGGACTGCATGTTCCGTGTCCTGTGGAGGAGGGATTCAGAGACGGAGCTT 1108

Qy       1329 TGTGTGTGTAG 1339
              |||||||||||
Db       1109 TGTGTGTGTAG 1119


RESULT 17
ABA06468
ID   ABA06468 standard; cDNA; 5769 BP.
XX
AC   ABA06468;
XX
DT   16 JAN 2002  (first entry)
XX
DE   Human cDNA SEQ ID NO: 134.
XX
KW   Human; gene therapy; neural disorder; immune system disorder;
KW   muscular disorder; reproductive disorder; gastrointestinal disorder;
KW   pulmonary disorder; cardiovascular disorder; renal disorder;
KW   proliferative disorder; inflammation; ss.
XX
OS   Homo sapiens
XX
PN   WO200154474 A2.
XX
PD   02 AUG 2001.
XX
PF   17 JAN 2001; 2001WO US01349.
XX
PR   31 JAN 2000; 2000US 179065P.
PR   04 FEB 2000; 2000US 180628P.
PR   24 FEB 2000; 2000US 184664P.
PR   02 MAR 2000; 2000US 186350P.
PR   16 MAR 2000; 2000US 189874P.
PR   17 MAR 2000; 2000US 190076P.
PR   18 APR 2000; 2000US 198123P.
PR   19 MAY 2000; 2000US 205515P.
PR   07 JUN 2000; 2000US 209467P.
PR   28 JUN 2000; 2000US 214886P.
PR   30 JUN 2000; 2000US 215135P.
PR   07 JUL 2000; 2000US 216647P.
PR   07 JUL 2000; 2000US 216880P.
PR   11 JUL 2000; 2000US 217487P.
PR   11 JUL 2000; 2000US 217496P.
PR   14 JUL 2000; 2000US 218290P.
PR   26 JUL 2000; 2000US 220963P.
PR   27 JUL 2000; 2000US 220964P.
PR   14 AUG 2000; 2000US 224518P.
PR   14 AUG 2000; 2000US 224519P.
PR   14 AUG 2000; 2000US 225213P.
PR   14 AUG 2000; 2000US 225214P.
PR   14 AUG 2000; 2000US 225266P.
PR   14-AUG-2000; 2000US-225267P.
PR   14-AUG-2000; 2000US-225268P.
PR   14-AUG-2000; 2000US-225270P.
PR   14-AUG-2000; 2000US-225447P.
PR   14-AUG-2000; 2000US-225757P.
PR   14-AUG-2000; 2000US-225758P.
PR   14-AUG-2000; 2000US-225759P.
PR   18-AUG-2000; 2000US-226279P.
PR   22-AUG-2000; 2000US-226681P.
PR   22-AUG-2000; 2000US-226868P.
PR   22-AUG-2000; 2000US-227182P.
PR   23-AUG-2000; 2000US-227009P.
PR   30-AUG-2000; 2000US-228924P.
PR   01-SEP-2000; 2000US-229287P.
PR   01-SEP-2000; 2000US-229343P.
PR   01-SEP-2000; 2000US-229344P.
PR   01-SEP-2000; 2000US-229345P.
PR   05-SEP-2000; 2000US-229509P.
PR   05-SEP-2000; 2000US-229513P.
PR   06-SEP-2000; 2000US-230437P.
PR   06-SEP-2000; 2000US-230438P.
PR   08-SEP-2000; 2000US-231242P.
PR   08-SEP-2000; 2000US-231243P.
PR   08-SEP-2000; 2000US-231244P.
PR   08-SEP-2000; 2000US-231413P.
PR   08-SEP-2000; 2000US-231414P.
PR   08-SEP-2000; 2000US-232080P.
PR   08-SEP-2000; 2000US-232081P.
PR   12-SEP-2000; 2000US-231968P.
PR   14-SEP-2000; 2000US-232397P.
PR   14-SEP-2000; 2000US-232398P.
PR   14-SEP-2000; 2000US-232399P.
PR   14-SEP-2000; 2000US-232400P.
PR   14-SEP-2000; 2000US-232401P.
PR   14-SEP-2000; 2000US-233063P.
PR   14-SEP-2000; 2000US-233064P.
PR   14-SEP-2000; 2000US-233065P.
PR   21-SEP-2000; 2000US-234223P.
PR   21-SEP-2000; 2000US-234274P.
PR   25-SEP-2000; 2000US-234997P.
PR   25-SEP-2000; 2000US-234998P.
PR   26-SEP-2000; 2000US-235484P.
PR   27-SEP-2000; 2000US-235834P.
PR   27-SEP-2000; 2000US-235836P.
PR   29-SEP-2000; 2000US-236327P.
PR   29-SEP-2000; 2000US-236367P.
PR   29-SEP-2000; 2000US-236368P.
PR   29-SEP-2000; 2000US-236369P.
PR   29-SEP-2000; 2000US-236370P.
PR   02-OCT-2000; 2000US-236802P.
PR   02-OCT-2000; 2000US-237037P.
PR   02-OCT-2000; 2000US-237038P.
PR   02-OCT-2000; 2000US-237039P.
PR   02-OCT-2000; 2000US-237040P.
PR   13-OCT-2000; 2000US-239935P.
PR   13-OCT-2000; 2000US-239937P.
PR   20-OCT-2000; 2000US-240960P.
PR   20-OCT-2000; 2000US-241221P.
PR   20-OCT-2000; 2000US-241785P.
PR   20-OCT-2000; 2000US-241786P.
PR   20-OCT-2000; 2000US-241787P.
PR   20-OCT-2000; 2000US-241808P.
PR   20-OCT-2000; 2000US-241809P.
PR   20-OCT-2000; 2000US-241826P.
PR   01-NOV-2000; 2000US-244617P.
PR   08-NOV-2000; 2000US-246474P.
PR   08-NOV-2000; 2000US-246475P.
PR   08-NOV-2000; 2000US-246476P.
PR   08-NOV-2000; 2000US-246477P.
PR   08-NOV-2000; 2000US-246478P.
PR   08-NOV-2000; 2000US-246523P.
PR   08-NOV-2000; 2000US-246524P.
PR   08-NOV-2000; 2000US-246525P.
PR   08-NOV-2000; 2000US-246526P.
PR   08-NOV-2000; 2000US-246527P.
PR   08-NOV-2000; 2000US-246528P.
PR   08-NOV-2000; 2000US-246532P.
PR   08-NOV-2000; 2000US-246609P.
PR   08-NOV-2000; 2000US-246610P.
PR   08-NOV-2000; 2000US-246611P.
PR   08-NOV-2000; 2000US-246613P.
PR   17-NOV-2000; 2000US-249207P.
PR   17-NOV-2000; 2000US-249208P.
PR   17-NOV-2000; 2000US-249209P.
PR   17-NOV-2000; 2000US-249210P.
PR   17-NOV-2000; 2000US-249211P.
PR   17-NOV-2000; 2000US-249212P.
PR   17-NOV-2000; 2000US-249213P.
PR   17-NOV-2000; 2000US-249214P.
PR   17-NOV-2000; 2000US-249215P.
PR   17-NOV-2000; 2000US-249216P.
PR   17-NOV-2000; 2000US-249217P.
PR   17-NOV-2000; 2000US-249218P.
PR   17-NOV-2000; 2000US-249244P.
PR   17-NOV-2000; 2000US-249245P.
PR   17-NOV-2000; 2000US-249264P.
PR   17-NOV-2000; 2000US-249265P.
PR   17-NOV-2000; 2000US-249297P.
PR   17-NOV-2000; 2000US-249299P.
PR   17-NOV-2000; 2000US-249300P.
PR   01-DEC-2000; 2000US-250160P.
PR   01-DEC-2000; 2000US-250391P.
PR   05-DEC-2000; 2000US-251030P.
PR   05-DEC-2000; 2000US-251988P.
PR   05-DEC-2000; 2000US-256719P.
PR   06-DEC-2000; 2000US-251479P.
PR   08-DEC-2000; 2000US-251856P.
PR   08-DEC-2000; 2000US-251868P.
PR   08-DEC-2000; 2000US-251869P.
PR   08-DEC-2000; 2000US-251989P.
PR   08-DEC-2000; 2000US-251990P.
PR   11-DEC-2000; 2000US-254097P.
PR   05-JAN-2001; 2001US-259678P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Barash SC,  Ruben SM;
XX
DR   WPI; 2001-476161/51.
DR   P-PSDB; ABB10246.
XX
PT   Isolated nucleic acid molecule encoding an inflammation-associated
PT   polypeptide is used in preventing, treating or ameliorating a medical
PT   condition  -
XX
PS   Claim 1; SEQ ID NO: 134; 859pp + Sequence Listing; English.
XX
CC   The present invention provides human cDNAs, proteins and related genomic
CC   DNAs. These can be used in the treatment of neural, immune system,
CC   muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,
CC   renal and proliferative disorders and inflammation. The present sequence
CC   is a cDNA of the invention.
XX
SQ   Sequence 5769 BP; 1299 A; 1695 C; 1628 G; 1146 T; 1 other;

  Query Match             18.2%;  Score 925.4;  DB 22;  Length 5769;
  Best Local Similarity   59.4%;  Pred. No. 1.7e-277;
  Matches 1701;  Conservative    1;  Mismatches 1067;  Indels   56;  Gaps    4;

Qy        186 CCAGACCTCAAGAAACACTCGTTCAGATGAAGACAAAGATGGCAACTGGGATGCTTGGGG 245
              || ||  || || | | | || || || || |||   || |||   |||||||  |||||
Db        124 CCTGAGTTCCAGGACCGCACGGCTCCGAGGAGGACCGGGACGGCCTATGGATGCCTGGGG 183

Qy        246 CGACTGGAGTGACTGCTCCCGGACCTGTGGGGAGGAGCCATCATATTCTCTGCGGAGATG 305
              |   |||||||| ||||| || ||||| ||||| || || || || |||||| || | ||
Db        184 CCCATGGAGTGAATGCTCACGGCACCTGCGGGGGTGGGCCTCCTACTCTCTGAGGCGCTG 243
```

```
Qy    306 [illegible] 365
Db    244 [illegible] 303

Qy    366 [illegible] 425
Db    304 [illegible] 363

Qy    426 [illegible] 485
Db    364 [illegible] 423

Qy    486 [illegible] 545
Db    424 [illegible] 483

Qy    546 [illegible] 605
Db    484 [illegible] 543

Qy    606 [illegible] 665
Db    544 [illegible] 603

Qy    666 [illegible] 725
Db    604 [illegible] 663

Qy    726 [illegible] 785
Db    664 [illegible] 723

Qy    786 [illegible] 845
Db    724 [illegible] 783

Qy    846 [illegible] 905
Db    784 [illegible] 843

Qy    906 [illegible] 965
Db    844 [illegible] 903

Qy    966 [illegible] 1025
Db    904 [illegible] 963

Qy   1026 [illegible] 1085
Db    964 [illegible] 1023

Qy   1086 [illegible] 1145
Db   1024 [illegible] 1083

Qy   1146 [illegible] 1205
Db   1084 [illegible] 1143

Qy   1206 [illegible] 1265
Db   1144 AGCCA [illegible] 1152

Qy   1266 [illegible] 1325
Db   1153 [illegible] 1212

Qy   1326 [illegible] 1385
Db   1213 [illegible] 1272

Qy   1386 [illegible] 1445
Db   1273 [illegible] 1332

Qy   1446 [illegible] 1505
Db   1333 [illegible] 1392

Qy   1506 [illegible] 1565
Db   1393 [illegible] 1452

Qy   1566 [illegible] 1625
Db   1453 [illegible] 1512

Qy   1626 [illegible] 1685
Db   1513 [illegible] 1572

Qy   1686 [illegible] 1745
Db   1573 [illegible] 1632

Qy   1746 [illegible] 1805
Db   1633 [illegible] 1692

Qy   1806 [illegible] GA 1857
Db   1693 [illegible] 1752

Qy   1858 [illegible] 1910
Db   1753 [illegible] 1812

Qy   1911 [illegible] 1970
Db   1813 [illegible] 1872

Qy   1971 [illegible] 2030
Db   1873 [illegible] 1932

Qy   2031 [illegible] 2090
Db   1933 [illegible] 1992

Qy   2091 [illegible] 2150
Db   1993 [illegible] 2052

Qy   2151 [illegible] 2197
Db   2053 [illegible] 2112

Qy   2198 [illegible] 2252
Db   2113 [illegible] 2172

Qy   2253 [illegible] 2312
Db   2173 [illegible] 2232

Qy   2313 [illegible] 2372
Db   2233 [illegible] 2292

Qy   2373 [illegible] 2432
Db   2293 [illegible] 2352

Qy   2433 [illegible] 2492
```

```
                |  | ||    |     ||||||||     | |    | |   ||||||| |  ||||
Db       2353 GCAAGCATGCAAGAAAGATGACTGTCCCAGCGAGTGGCTTCTCTCAGACTGGACAGACTC 2412

Qy       2493 TTCTGTCAGTTGTGGTGTTGGAATCCAGAGAAGAAAGCAGGTGTGTCAAAGGCTGGCAGC 2552
              |||    || || || |  || | |||||   |||      | || | || | ||
Db       2413 TTCCACAAGCTGCGGGGAAGGCACCCAGACTCGAAGCGCCATTTGCCGAAAGATGCTGAA 2472

Qy       2553 CAAAGGTCGGCGCATCCCCCTCAGTGAGATGATGTGCAGGGATCTACCAGGGTTCCCTCT 2612
               |  || |     |      ||| |   |   |||||  |   || ||    |   |
Db       2473 AACCGGCCTCTCAACGGTTGTCAATTCCACCCTGTGCCCGCCCCTGCCTTTCTCTTCCTC 2532

Qy       2613 TGTAAGATCTTGCCAGATGCCTGAGTGCAGTAAAATCAAATCAGAGATGAAGACAAAACT 2672
                | ||    ||   |  |||            || |    ||  |     |     |
Db       2533 CATCAGGCCCTGTATGCTGGC------------AACCTGTGCAAGGCCCGGGCGGCCATC 2580

Qy       2673 TGGTGAGCAGGGTCGGCAGATCCTCAGTGTCCAGAGAGTCTACATTCAGACAAGGGAAGA 2732
                   |||| | ||||| |||      | |  ||  |||||||| |||||  |     |
Db       2581 CACGAAGCACAGCCCGCACATCGCGGCCGCCAGGAAGGTCTACATACAGACTCGCAGGCA 2640

Qy       2733 GAAGCGTATTAACCTGACCATTGGTAGCAGAGCCTATTTGCTGCCCAACACATCCGTGAT 2792
              || |    |  || |     | ||  ||   |||||  |||| ||||| ||  | ||| |
Db       2641 GAGGAAGCTGCACTTCGTGGTGGGGGGCTTCGCCTACCTGCTCCCCAAGACGGCGGTGGT 2700

Qy       2793 TATTAAGTGCCCCGTGCGACGATTCCAGAAATCTCTCATCCAGTGGGAGAAGGATGGCCG 2852
                |    ||||| | |||  |  |||  ||  | || |||   ||||||||||| ||||
Db       2701 GCTGCGCTGCCCGGCGGCGGCAGGTCCGCAAGCCCCTCATCACCTGGAGAAGGACGGCCA 2760

Qy       2853 TTGCCTGCAGAACTCCAAACGGCTTGGCATCACCAAGTCAGGCTCACTAAAAATCCACGG 2912
                 |||    | ||| |  |   |     |  ||   |  ||||  || || |||||| |
Db       2761 GCACCTCATCAGCTCGACGCACGTCACGGTGGCCCCCTTCGGCTATCTCAAGATCCACCG 2820

Qy       2913 TCTTGCTGCCCCCGACATCGGCGTGTACCGGTGCATTGCAGGCTCTGCACAGGAAACAGT 2972
               ||     || | ||    ||||| |||   |||   || ||| | || | |||     |
Db       2821 CCTCAAGCCCTCGGATGCAGGCGTCTACACCTGCTCAGCGGGCCCGGCCCGGGAGCACTT 2880

Qy       2973 TGTGCTCAAGCTCATTGGTACTGACAACCGGCTCATCGCACGCCC 3017
              |||| | |||||||| ||     ||  |  |||| | || || ||
Db       2881 TGTGATTAAGCTCATCGGAGGCAACCGCAAGCTCGTGGCCCGGCC 2925

Search completed: March 22, 2003, 06:47:35
Job time : 860 secs
```

```
                    GenCore version 5.1.4_p5_4578
           Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:          March 22, 2003, 06:01:08 ; Search time 4976 Seconds
                                             (without alignments)
                                             16540.904 Million cell updates/sec

Title:           US-09-784-358-1
Perfect score:   5076
Sequence:        1 atggcttcctggacgagccc..........agtcatgtcaaaaggataa 5076

Scoring table:   IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0

Searched:        16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters:      32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries

Database :        EST:*
                 1:  em_estba:*
                 2:  em_esthum:*
                 3:  em_estin:*
                 4:  em_estmu:*
                 5:  em_estov:*
                 6:  em_estpl:*
                 7:  em_estro:*
                 8:  em_htc:*
                 9:  gb_est1:*
                 10:  gb_est2:*
                 11:  gb_htc:*
                 12:  gb_est3:*
                 13:  gb_est4:*
                 14:  gb_est5:*
                 15:  em_estfun:*
                 16:  em_estom:*
                 17:  gb_gss:*
                 18:  em_gss_hum:*
                 19:  em_gss_inv:*
                 20:  em_gss_pln:*
                 21:  em_gss_vrt:*
                 22:  em_gss_fun:*
                 23:  em_gss_mam:*
                 24:  em_gss_mus:*
                 25:  em_gss_other:*
                 26:  em_gss_pro:*
                 27:  em_gss_rod:*

    Pred. No. is the number of results predicted by chance to have a
    score greater than or equal to the score of the result being printed,
    and is derived by analysis of the total score distribution.

                               SUMMARIES
                %
Result          Query
  No.   Score   Match Length DB  ID                        Description

    1   603.4   11.9    606  12  BG750735                  BG750735 602706924
c   2     474    9.3    485   9  AA707140                 AA707140 zj36e01.s
    3   409.6    8.1    846  11  AK020361                  AK020361 Mus muscu
    4   401.4    7.9    425  12  BF767952                  BF767952 CM1-CN006
    5   394.2    7.8   1009  13  BM476141                  BM476141 AGENCOURT
c   6   293.6    5.8    338   9  AA906029                 AA906029 oj90a07.s
```

```
    7   293.6    5.8    622  10  BB633755                  BB633755 BB633755
    8   281.8    5.6    890  17  AQ743958                  AQ743958 HS_5504_A
    9     278    5.5    850  12  BE888902                  BE888902 601513828
   10     274    5.4    619  10  BB478282                  BB478282 BB478282
c  11   267.8    5.3    271  10  BE162337                  BE162337 PM2-HT045
   12   266.8    5.3    979  11  AK020115                  AK020115 Mus muscu
   13   262.4    5.2    660  10  BB659292                  BB659292 BB659292
c  14   259.6    5.1    458  12  BF543337                  BF543337 UI-R-Y0-a
   15     256    5.0    581  13  BI523618                  BI523618 603051752
   16     256    5.0    755  10  BB643318                  BB643318 BB643318
   17   252.2    5.0    439  10  AW631803                  AW631803 91162 MAR
   18   224.8    4.4    591  13  BM088969                  BM088969 502594 MA
c  19   224.8    4.4    782  14  BQ007587                  BQ007587 UI-H-EI1
c  20   214.6    4.2    644  12  BF111214                  BF111214 7n44e07.x
   21   214.2    4.2    559  10  BE628435                  BE628435 uu09f09.y
c  22   207.8    4.1    608  10  BE631939                  BE631939 uu09f09.x
   23   207.6    4.0    423   9  AA482392                 AA482392 zv05h07.r
c  24   197.6    3.9    631  12  BF445136                  BF445136 nad21a04.
c  25   183.8    3.6    586  12  BF510853                  BF510853 UI-H-BI4
c  26   182.8    3.6    712  17  AZ939026                  AZ939026 2M0197N18
c  27   179.8    3.5    628   9  AI459225                 AI459225 tk11b03.x
c  28   166.2    3.3    452  10  AW120794                  AW120794 UI-M-BH2
   29   159.2    3.1    419  14  H26489                    H26489 EST33E01 BA
   30   148.4    2.9    468  17  AZ883573                  AZ883573 RPCI-23-1
   31   148.4    2.9    553  17  AZ882856                  AZ882856 RPCI-23-1
   32     146    2.9    637  10  BB612331                  BB612331 BB612331
   33   140.6    2.8    433  14  W47220                    W47220 zc39a07.r1
   34   140.6    2.8    442  10  BB839098                  BB839098 BB839098
   35     135    2.7    553  13  BI472859                  BI472859 ft02g02.y
c  36   132.8    2.6    534   9  AA971699                 AA971699 op05a11.s
c  37   128.2    2.5    253  12  BF848902                  BF848902 QV1-EN0
   38   121.6    2.4    302   9  AA284278                 AA284278 zt39a07.r1
c  39   121.2    2.4    289  10  BE162336                  BE162336 PM2-HT045
   40   109.6    2.2    366  14  BQ561231                  BQ561231 H4006A11
c  41   109.6    2.2    437  14  W47029                    W47029 zc39a07.s1
c  42   108.2    2.1    568   9  AI917724                 AI917724 tt11c08.x
c  43   107.2    2.1    522   9  AI342006                 AI342006 qs08d10.x
   44   100.6    2.0    657  17  AZ043795                  AZ043795 RPCI-23-3
   45    98.2    1.9    460  17  AZ408665                  AZ408665 1M0281N20


                              ALIGNMENTS


RESULT 1
BG750735
LOCUS       BG750735                 606 bp    mRNA    linear   EST 15-MAY-2001
DEFINITION  602706924F1 NIH_MGC_43 Homo sapiens cDNA clone IMAGE:4843636 5',
            mRNA sequence.
ACCESSION   BG750735
VERSION     BG750735.1  GI:14061388
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 606)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Lina Hood/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1679  row: p  column: 05
            High quality sequence stop: 606.
FEATURES             Location/Qualifiers
     source          1..606
```

```
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4843636"
                     /clone_lib="NIH_MGC_43"
                     /tissue_type "normal pigmented retinal epithelium"
                     /lab_host="DH10B (phage-resistant)"
                     /note "Organ: eye; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI, cDNA made by oligo dT priming.  Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG(G).  Library constructed  by Ling Hong
                     in the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using ZAP-cDNA synthesis kit
                     (Stratagene) and  Superscript II RT (Life Technologies).
                     Note: this is a NIH_MGC Library. |"
BASE COUNT      172 a    135 c    144 g    155 t
ORIGIN

  Query Match             11.9%;  Score 603.4;  DB 12;  Length 606;
  Best Local Similarity   99.8%;  Pred. No. 2.5e-157;
  Matches  604;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy     732 AGAAAATGTAATTGCTGTTCCTTTGGGAAGTCGAAGTGTGAGAATTACAGTGAAAGGACC 791
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2 AGAAAATGTAATTGCTGTTCCTTTGGGAAGTCGAAGTGTGAGAATTACAGTGAAAGGACC 61

Qy     792 TGCCCACCTCTTTATTGAATCAAAAACACTTCAAGGAAGCAAAGGAGAACACAGCTTTAA 851
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      62 TGCCCACCTCTTTATTGAATCAAAAACACTTCAAGGAAGCAAAGGAGAACACAGCTTTAA 121

Qy     852 CAGCCCCGGCGTCTTTGTCGTAGAAAACACAACAGTGGAATTTCAGAGGGCCTCCCACAC 911
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     122 CAGCCCCGGCGTCTTTGTCGTAGAAAACACAACAGTGGAATTTCAGAGGGCCTCCCACAC 181

Qy     912 GCAAACTTTTAAGATTCCAGGACCTCTGATGGCTGATTTCATCTTCAAGACCAGGTACAC 971
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     182 GCAAACTTTTAAGATTCCAGGACCTCTGATGGCTGATTTCATCTTCAAGACCAGGTACAC 241

Qy     972 TGCAGCCAAAGACAGCGTGGTTCAGTTCTTCTTTTACCAGCCCATCAGTCATCAGTGGAG 1031
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     242 TGCAGCCAAAGACAGCGTGGTTCAGTTCTTCTTTTACCAGCCCATCAGTCATCAGTGGAG 301

Qy    1032 ACAAACTGACTTCTTTCCCTGCACTGTGACGTGTGGAGGAGGTTATCAGCTCAATTCTGC 1091
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     302 ACAAACTGACTTCTTTCCCTGCACTGTGACGTGTGGAGGAGGTTATCAGCTCAATTCTGC 361

Qy    1092 TGAATGTGTGGATATCCGCTTGAAGAGGGTAGTTCCTGACCATTATTGTCACTACTACCC 1151
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     362 TGAATGTGTGGATATCCGCTTGAAGAGGGTAGTTCCTGACCATTATTGTCACTACTACCC 421

Qy    1152 TGAAAATGTAAAACCAAAACCAAAACTGAAGGAATGCAGCATGGATCCCTGCCCATCAAG 1211
           ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db     422 TGAAAATGTAAAACCAAAAGCAAAACTGAAGGAATGCAGCATGGATCCCTGCCCATCAAG 481

Qy    1212 TGATGGATTTAAAGAGATAATGCCCTATGACCACTTCCAACCTCTTCCTCGCTGGGAACA 1271
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     482 TGATGGATTTAAAGAGATAATGCCCTATGACCACTTCCAACCTCTTCCTCGCTGGGAACA 541

Qy    1272 TAATCCTTGGACTGCATGTTCCGTGTCCTGTGGAGGAGGGATTCAGAGACGGAGCTTTGT 1331
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     542 TAATCCTTGGACTGCATGTTCCGTGTCCTGTGGAGGAGGGATTCAGAGACGGAGCTTTGT 601

Qy    1332 GTGTG 1336
           |||||
Db     602 GTGTG 606


RESULT 2
AA707140/c
LOCUS       AA707140                 485 bp    mRNA    linear   EST 24-DEC-1997
DEFINITION  zj33e01.s1 Soares_fetal_liver_spleen_1NFLS_S1 Homo sapiens cDNA
            clone IMAGE:452088 3' similar to TR:P97857 P97857 SECRETORY PROTEIN
            CONTAINING THROMBOSPONDIN MOTIFS. ;, mRNA sequence.
ACCESSION   AA707140
```

```
VERSION     AA707140.1  GI:2717058
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 485)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
            ,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
            White,Y., Wylie,T., Waterston,R. and Wilson,R.
  TITLE     WashU-NCI human EST Project
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 471.
FEATURES             Location/Qualifiers
     source          1..485
                     /organism="Homo sapiens"
                     /db_xref="GDB:1388444"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:452088"
                     /clone_lib="Soares_fetal_liver_spleen_1NFLS_S1"
                     /sex="male"
                     /dev_stage="20 week-post conception fetus"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
                     with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
                     This is a subtracted version of the original Soares fetal
                     liver spleen 1NFLS library.  1st strand cDNA was primed
                     with a Pac I - oligo(dT) primer [5'
                     AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Pac I and cloned into the Pac I
                     and Eco RI sites of the modified pT7T3 vector.  Library
                     went through one round of normalization. Library
                     constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT      108 a    144 c    114 g    119 t
ORIGIN

  Query Match              9.3%;  Score 474;  DB 9;  Length 485;
  Best Local Similarity   99.8%;  Pred. No. 4.2e-121;
  Matches  485;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy     236 ATGCTTGGGGCGACTGGAGTGACTGCTCCCGGACCTGTGGGGGAGGAGCATCATATTCTC 295
           |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db     485 ATGCTTGGGGCGACTGGAGTGACTGCTCCCGGACCTGT GGGGAGGAGCATCATATTCTC 427

Qy     296 TGCGGAGATGTTTGACTGGAAGGAATTGTGAAGGGCAGAACATTCGGTACAAGACATGCA 355
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     426 TGCGGAGATGTTTGACTGGAAGGAATTGTGAAGGGCAGAACATTCGGTACAAGACATGCA 367

Qy     356 GCAATCATGACTGCCCTCCAGATGCAGAAGATTTCAGAGCCCAGCAGTGCTCAGCCTACA 415
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     366 GCAATCATGACTGCCCTCCAGATGCAGAAGATTTCAGAGCCCAGCAGTGCTCAGCCTACA 307

Qy     416 ATGATGTCCAGTATCAGGGGCATTACTATGAATGGCTTCCACGATATAATGATCCTGCTG 475
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     306 ATGATGTCCAGTATCAGGGGCATTACTATGAATGGCTTCCACGATATAATGATCCTGCTG 247

Qy     476 CCCCGTGTGCACTCAAGTGTCATGCACAAGGACAAAACTTGGTGGTGGAGCTGGCACCTA 535
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     246 CCCCGTGTGCACTCAAGTGTCATGCACAAGGACAAAACTTGGTGGTGGAGCTGGCACCTA 187

Qy     536 AGGTACTGGATGGAACTCGTTGCAACACGGACTCCTTGGACATGTGTATCAGTGGCATCT 595
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 A[illegible] 127

Qy        596 [illegible] 655
              [illegible]
Db        126 [illegible] 67

Qy        656 TCTGTGCCGGCGATGGCTCCACTGCAGGCTTGTACGGGACAATCAAAGTCACACGTTT 715
              [illegible]
Db         66 TCTGTGCCGGCGATGGCTCCACTGCAGGCTTGTACGGGACAATCAAAGTCACACGTTT 7

Qy        716 CTCCTG 721
              ||||||
Db          6 CTCCTG 1


RESULT 3
AK020361
LOCUS       AK020361                 846 bp    mRNA    linear   HTC 19-JAN-2002
DEFINITION  Mus musculus adult male epididymis cDNA, RIKEN full-length enriched
            library, clone:9230119C12:thrombospondin type 1 domain containing
            protein, full insert sequence.
ACCESSION   AK020361
VERSION     AK020361.1  GI:12[illegible]
KEYWORDS    HTC; CAP trapper.
SOURCE      Mus musculus (strain:C57BL/6J) adult male epididymis cDNA to mRNA,
            clone_lib:RIKEN full-length enriched mouse cDNA library
            clone:9230119C12.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Carninci,P. and Hayashizaki,Y.
  TITLE     High-efficiency full-length cDNA cloning
  JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
  MEDLINE   99279253
   PUBMED   10349636
REFERENCE   2
  AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
  TITLE     Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
  JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
  MEDLINE   20499374
   PUBMED   11042159
REFERENCE   3
  AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
            Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
            Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
            Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
            Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
            Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
            Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
  TITLE     RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer
  JOURNAL   Genome Res. 10 (11), 1757-1771 (2000)
  MEDLINE   20530913
   PUBMED   11076861
REFERENCE   4
  AUTHORS   Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y.,
            Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S.,
            Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I.,
            Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R.,
            Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T.,
            Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H.,
            Kuehl,P., Lewis,S., Matsuo,Y., Nikaido,I., Pesole,G.,
            Quackenbush,J., Schriml,L.M., Staubli,F., Suzuki,R., Tomita,M.,
            Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H.,
            Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N.,
            Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C.,
            Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D.,
            Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P.,
            Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P.,
            Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H.,
            Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H.,
            Toyo-oka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L.,
            Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohtsuki,S.
            and Hayashizaki,Y.
  TITLE     Functional annotation of a full-length mouse cDNA collection
  JOURNAL   Nature 409 (6821), 685-690 (2001)
  MEDLINE   21085660
   PUBMED   11217851
REFERENCE   5  (bases 1 to 846)
  AUTHORS   Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A.,
            Arakawa,T., Baldarelli,R., Bono,H., Brownstein,M., Bult,C.,
            Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T.,
            Hara,A., Hayatsu,N., Hill,D., Hiramoto,K., Hiraoka,T., Hori,F.,
            Hume,D., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T.,
            Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S.,
            Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K.,
            Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Quackenbush,J.,
            Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D.,
            Schriml,L., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T.,
            Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,
            Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yamanaka,I.,
            Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and
            Hayashizaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-AUG-2000) Yoshihide Hayashizaki, The Institute of
            Physical and Chemical Research (RIKEN), Laboratory for Genome
            Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
            RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
            Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
            URL:http://genome-res.gsc.riken.go.jp/, Tel:81-45-503-9222,
            Fax:81-45-503-9216)
COMMENT     Please visit our web site (http://genome.gsc.riken.go.jp/) for
            further details.
            cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues. First strand cDNA was primed with a primer
            [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTTTVN 3'], cDNA was
            prepared by using trehalose thermo-activated reverse transcriptase
            and subsequently enriched for full-length by cap-trapper. cDNA went
            through one round of normalization to Rot = 10.0 and subtraction to
            Rot = 185.2. Second strand cDNA was prepared with the primer
            adapter of sequence [5'
            GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCC 3'], cDNA was cleaved
            with BamHI and XhoI. Vector: a modified pBluescript KS(+) after
            bulk excision from Lambda FLC-1. Cloning sites, 5' end: SalI; 3'
            end: BamHI. Host: DH10B.
FEATURES             Location/Qualifiers
     source          1..846
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="FANTOM_DB:9230119C12"
                     /db_xref="MGD:MGI:1911798"
                     /db_xref="taxon:10090"
                     /clone="9230119C12"
                     /sex="male"
                     /tissue_type="epididymis"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="adult"
     misc_feature    1..846
                     /note="Thrombospondin type 1 domain containing protein
                     data source:Pfam, source key:PF00090, evidence:ISS"
                     /db_xref="MGD:MGI:1925159"
BASE COUNT      227 a    202 c    219 g    198 t
ORIGIN

  Query Match             8.1%;  Score 4[illegible];  DB 11;  Length 846;
  Best Local Similarity   84.6%;  Pred. No. [illegible];
  Matches  460;  Conservative    0;  Mismatches   84;  Indels    0;  Gaps    0;

Qy        185 [illegible] 244
              [illegible]
```

```
Db     150 ACAAGACTTCAAGAAATGCTCATTCAGAAGAAGACAAGGATGGCAACTGGGACGCCTGGG 209

Qy     245 GCGACTGGAGTGACTGCTCCCGGACCTGTGGCGGAGCAGCATCATATTCTCTCCCGAGAT 304
           | ||||||||||||||||||||||||||||| |||||||| || |||||| ||||||||
Db     210 GAGACTGGAGTGACTGCTCCCGGACCTGTGGCGGAGGAGCCTCCTATTCTTTGCGGAGAT 269

Qy     305 GTTTGACTGGAAGGAATTGTGAAGGGCAGAACATTCGGTACAAGACATGCAGCAATCATG 364
           || ||||||||||||| || ||||| ||||||||   | |||||||| |||||||||||| |
Db     270 GTCTGACTGGAAGGAACTGCGAAGGACAGAACATCACATACAAGACTTGCAGCAATCACG 329

Qy     365 ACTGCCCTCCAGATGCAGAAGATTTCAGAGCCCAGCAGTGCTCAGCCTACAATCATCTCC 424
           |||||||   ||||||||||||| |||||||||||||||||||||||||||||||| || |
Db     330 ACTGCCCAGCAGATGCAGAAGACTTCAGAGCCCAGCAGTCCTCAGCCTACAATGACCTGC 389

Qy     425 AGTATCAGGGGCATTACTATGAATGGCTTCCACGATATAATGATCCTGCTGCCCCCTGTG 484
           |||| ||||| ||||| |||||||||| |||||| ||| ||||| ||||||||||| ||||
Db     390 AGTACCAGGGTCATTATTATGAATGGATTCCACTATACAATGACCCTGCTGCCCCATGTG 449

Qy     485 CACTCAAGTGTCATGCACAAGGACAAAACTTGGTGGTGGAGCTGGCACCTAAGGTACTGG 544
           |||||||||||||| || |   ||||| |  |||||| | |||||||||||| || ||  |||
Db     450 CACTCAAGTGTCACGCTCGGGGACAGAGTTTGGTGGTAGAGCTGGCACCCAAAGTGTTGG 509

Qy     545 ATGGAACTCGTTGCAACACGGACTCCTTGGACATGTGTATCAGTGGCATCTGTCAGGCAG 604
           ||:||||||| |||||   ! |||||   |||| ||| | || ||||||||||| |||    |
Db     510 ATGGAACTCGCTGCAATGCTGACTCTCTGGATATGTGCATAAGTGGCATCTGCCAGATTG 569

Qy     605 TGGGCTGCGATCGGCAACTGGGAAGCAATGCCAAGGAGGACAACTGTGGAGTCTGTGCCG 664
           ||||||| ||| | || |||||||| |||||||||:||||||| ||||| |||||||| |
Db     570 TGGGCTGTGATAGACAGCTGGGAAGTAATGCCAAGGAGGACAATTGTGGCGTCTGTGCTG 629

Qy     665 GCGATGGCTCCACCTGCAGGCTTGTACGGGGACAATCAAAGTCACACGTTTCTCCTGAAA 724
           : || ||:|:||||:'||||: || ||||| ||  | |||   ||| : |||||||| |
Db     630 GGGACGGCTCCACCTGCAGGCTGGTGCGGGGTCAGACCAAGATGCACCTGTCTCCTGAGA 689

Qy     725 AAAG 728
           | ||
Db     690 AGAG 693
```

```
RESULT 4
BF767952
LOCUS       BF767952                 425 bp    mRNA    linear   EST 12-JAN 2001
DEFINITION  CM1-CN0061-201200-673 h01 CN0061 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BF767952
VERSION     BF767952.1  GI:12115852
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 425)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A , Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=CM1&t2=CM1-CN0061
            201200-673-h01&t3=2000-12-20&t4=1)
```

```
            Seq primer: puc 18 forward
            High quality sequence start: 12
            High quality sequence stop: 425.
FEATURES             Location/Qualifiers
     source          1..425
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="CN0061"
                     /dev_stage="Adult"
                     /note="Organ: colon_normal; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716    Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
BASE COUNT      112 a    100 c     98 g    115 t
ORIGIN

  Query Match             7.9%;  Score 401.4;  DB 12;  Length 425;
  Best Local Similarity   99.5%;  Pred. No. 8.7e-101;
  Matches  413;  Conservative    0;  Mismatches    1;  Indels    1;  Gaps    1;

Qy     915 AACTTTTAAGATTCCAGGACCTCTGATGGCTGATTTCATCTTCAAGACCAGGTACACTGC 974
           |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db      12 AACTTTTAAGATTCCAGGACCTCTGATGGC-GATTTCATCTTCAAGACCAGGTACACTGC 70

Qy     975 AGCCAAAGACAGCGTGGTTCAGTTCTTCTTTTACCAGCCCATCAGTCATCAGTGGAGACA 1034
           |||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db      71 AGCCAAAGACAGCGTGGTTCAGTTCTTCTTTTATCAGCCCATCAGTCATCAGTGGAGACA 130

Qy    1035 AACTGACTTCTTTCCCTGCACTGTGACGTGTGGAGGAGGTTATCAGCTCAATTCTGCTGA 1094
           |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db     131 AACTGACTTCTTTCCCTGCACTGTGACGTGTGGAGGAGGTTATCAGCTCAATTCTGCTGA 190

Qy    1095 ATGTGTGGATATCCGCTTGAAGAGGGTAGTTCCTGACCATTATTGTCACTACTACCCTGA 1154
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     191 ATGTGTGGATATCCGCTTGAAGAGGGTAGTTCCTGACCATTATTGTCACTACTACCCTGA 250

Qy    1155 AAATGTAAAACCAAAACCAAAACTGAAGGAATGCAGCATGGATCCCTGCCCATCAAGTGA 1214
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     251 AAATGTAAAACCAAAACCAAAACTGAAGGAATGCAGCATGGATCCCTGCCCATCAAGTGA 310

Qy    1215 TGGATTTAAAGAGATAATGCCCTATGACCACTTCCAACCTCTTCCTCGCTGGGAACATAA 1274
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     311 TGGATTTAAAGAGATAATGCCCTATGACCACTTCCAACCTCTTCCTCGCTGGGAACATAA 370

Qy    1275 TCCTTGGACTGCATGTTCCGTGTCCTGTGGAGGAGGGATTCAGAGACGGAGCTTT 1329
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     371 TCCTTGGACTGCATGTTCCGTGTCCTGTGGAGGAGGGATTCAGAGACGGAGCTTT 425
```

```
RESULT 5
BM476141
LOCUS       BM476141                1009 bp    mRNA    linear   EST 05-FEB-2002
DEFINITION  AGENCOURT_6478885 NIH_MGC_88 Homo sapiens cDNA clone IMAGE:5558669
            5', mRNA sequence.
ACCESSION   BM476141
VERSION     BM476141.1  GI:18525183
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1009)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation. Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
```

```
                     DNA Sequencing by: Agencourt Bioscience Corporation
                     Clone distribution: MGC clone distribution information can be
                     found through the I.M.A.G.E. Consortium/LLNL at:
                     http://image.llnl.gov
                     Plate: LLAM12282  row: a  column: 06
                     High quality sequence stop: 647.
FEATURES             Location/Qualifiers
     source          1..1999
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5558669"
                     /clone_lib="NIH_MGC_88"
                     /tissue_type="duodenal adenocarcinoma, cell line"
                     /lab_host="DH10B (phage resistant)"
                     /note="Organ: small intestine; Vector: pCMV-SPORT6;
                     Site 1: NotI; Site 2: SalI; Cloned unidirectionally;
                     oligo-dT primed. Average insert size 1.767 kb. Library
                     enriched for full length clones and constructed by Life
                     Technologies. Note: this is a NIH_MGC Library."
BASE COUNT      234 a    299 c    268 g    204 t      4 others
ORIGIN

  Query Match             7.8%;  Score 394.2;  DB 13;  Length 1009;
  Best Local Similarity   68.2%;  Pred. No. 3.1e-99;
  Matches  546;  Conservative    0;  Mismatches  255;  Indels    0;  Gaps    0;

Qy        932 GACCTCTGATGGCTGATTTCATCTTCAAGACCAGGTACACTGCAGCCAAAGACAGCGTGG 991
              |||| || |  || |||||||  ||||||    |  || | |    ||   |||||    |
Db         24 GACCACTCACAGCAGATTTCATTGTCAAGATTCGTAACTCGGGCTCCGCTGACAGTACAG 83

Qy        992 TTCAGTTCTTCTTTTACCAGCCCATCAGTCATCAGTGGAGACAAACTGACTTCTTTCCCT 1051
              | ||||| |||| || || |||||||   || |   |||||   | || || ||||| || |
Db         84 TCCAGTTCATCTTCTATCAACCCATCATCCACCGATGGAGGGAGACGGATTTCTTTCCTT 143

Qy       1052 GCACTGTGACGTGTGGAGGAGGTTATCAGCTCAATTCTGCTGAATGTGTGGATATCCGCT 1111
              || | |   || |||||||||||||||||||| |   || ||||| ||    ||| |  |
Db        144 GCTCAGCAACCTGTGGAGGAGGTTATCAGCTGACATCGGCTGAGTGCTACGATCTGAGGA 203

Qy       1112 TGAAGAGGGTAGTTCCTGACCCATTATTGTCACTACTACCCTGAAAATGTAAAACCAAAAC 1171
                 ||  | || ||| ||||||| || |||||||| ||||| || ||   | |||| ||||
Db        204 GCAACCGTGTGGTTGCTGACCAATACTGTCACTATTACCCAGAGAACATCAAACCCAAAC 263

Qy       1172 CAAAACTGAAGGAATGCAGCATGGATCCCTGCCCATCAAGTGATGGATTTAAAGAGATAA 1231
              | || ||   |||| |||| | ||||||| || ||| | |||| |||   ||   |||| |
Db        264 CCAAGCTTCAGGAGTGCAACTTGGATCCTTGTCCAGCCAGTGACGGATACAAGCAGATCA 323

Qy       1232 TGCCCTATGACCACTTCCAACCTCTTCCTCGCTGGGAACATAATCCTTGGACTGCATGTT 1291
              |||| |||||||| || ||| || |||||||| |||||        |  || |||| || || |
Db        324 TGCCTTATGACCTCTACCATCCCCTTCCTCGGTGGGAGGCCACCCCATGGACCGCGTGCT 383

Qy       1292 CCGTGTCCTGTGGAGGAGGGATTCAGAGACGGAGCTTTGTGTGTGTAGAGGAATCCATCC 1351
              ||    || ||||  | || || ||||| |||      ||    ||||| |||||   ||| |
Db        384 CCTCCTCGTGTGGGGGGGGCATCCAGAGCCGGGCAGTTTCCTGTGTGGAGGAGGACATCC 443

Qy       1352 ATGGAGAGATACTGCACCCAGAAGAATGGAAGTGCATGTACCCACCCAAACCCAAGCTTA 1411
              | ||   |   |           ||||||||| ||||| ||||||||| | || ||        |
Db        444 AGGGGCATGTCACTTCAGTGGAAGAGTGGAAATGCATGTACACCCCTAAGATGCCCATCG 503

Qy       1412 TGCAAACTTGTAATCTGTTTGATTGCCCAAGTGGATTGCCATGGAGTGGTCTCAGTGCA 1471
               |||   | || ||     ||||| ||||| || ||| | ||    |||||||||| |||||
Db        504 CGCAGCCCTGCAACACTTTTGACTGCCCTAAATGGCTGGCACAGGAGTGGTCTCCGTGCA 563

Qy       1472 CAGTGACTTGTGGCCGAGGGTTACGGTACCGGGTTGTTCTGTGTATTAACCACCGGGGAG 1531
              ||||||| ||||||||   ||    |   | ||||| || || || || ||   |||| || |||
Db        564 CAGTGACATGTGGCCAGGGCCTCAGATACCGTGTGGTCCTCTGCATCGACCATCGAGGAA 623

Qy       1532 AGCATGTTGGGGGCTGCAATCCACAACTGAAGTTACACATCAAAGAAGAATGTGTCATTC 1591
               |||      || ||||| |   ||| ||     |||     ||||| ||||| |||||   || | |
Db        624 TGCACACAGGAGGCTGTAGCCAAAAACAAAGCCCACATAAAAGAGGAATGCATCGTA 683

Qy       1592 CCATCCCGTGTTATAAACCAAAAGAAAAAAGTCCAGTGGAAGCAAAATTGCCTTGGCTGA 1651
              |||   || || |||||||||| ||||| |||   |||||| || || || ||||| ||| |
Db        684 CCACTCCCTGCTATAAACCCAAAGAGAAACTTCCAGTCGAGGCCAAGTTGCCATGGTTCA 743
```

```
Qy       1652 AACAAGCACAAGAACTAGAAGAGACCAGAATAGCAACAGAAGAACCAACGTTCATTCCAG 1711
              ||||||| ||||| |||||||             |||| ||     ||||||| | ||
Db        744 AACAAGCTCAAGAGCTAGAAGAAGGAGCTGCTGTGTCAGAGGAGCCCTCGTTCATCCCAG 803

Qy       1712 AACCCTGGTCAGCCTGCAGTA 1732
                 ||||||| | |||||
Db        804 AGGCCTGGTCGGNCTGCACNA 824


RESULT 6
AA906029/c
LOCUS       AA906029                 338 bp    mRNA    linear   EST 19-MAY-1998
DEFINITION  oj90a07.s1 Soares NFL_T_GBC_S1 Homo sapiens cDNA clone
            IMAGE:1505556 3', mRNA sequence.
ACCESSION   AA906029
VERSION     AA906029.1  GI:3041152
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 338)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 538   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 273.
FEATURES             Location/Qualifiers
     source          1..338
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1505556"
                     /clone_lib="Soares_NFL_T_GBC_S1"
                     /lab_host="DH10B"
                     /note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site 1: Not I; Site 2: Eco RI;
                     Equal amounts of plasmid DNA from three normalized
                     libraries (fetal lung NbHL19W, testis NHT, and B cell
                     NCI_CGAP_GCB1) were mixed, and ss circles were made in
                     vitro. Following HAP purification, this DNA was used as
                     tracer in a subtractive hybridization reaction. The driver
                     was PCR amplified cDNAs from pools of 5,000 clones made
                     from the same 3 libraries. The pools consisted of
                     I.M.A.G.E. clones 297480-302087, 682632-687239,
                     726408-728711, and 729096-731399. Subtraction by Bento
                     Soares and M. Fatima Bonaldo."
BASE COUNT       73 a     68 c     76 g    121 t
ORIGIN

  Query Match             5.8%;  Score 293.6;  DB 9;  Length 338;
  Best Local Similarity   98.7%;  Pred. No. 1.3e-70;
  Matches  296;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       3417 GGGCATCCAGGAAGAGACACCTCCTGCTGCTTAGCTCAGAGGGGAAACAGGGAGTGTGTT 3476
              ||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db        338 GGGCATCCAGGAAGAGACACCTCCTGCTGCTCAGCTCAGAGGGGAAACAGGGAGTGTGTC 279

Qy       3477 CCAAAGCTCGCATGCAAAAAACTCAGGCAAGCTGACATTCAAGCCGAAAGGACCTGTTCT 3536
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        278 CCAAAGCTCGCATGCAAAAAACTCAGGCAAGCTGACATTCAAGCCGAAAGGACCTGTTCT 219

Qy       3537 CATGAGGCAAAGCCAACCTCCTCAATTTCATTTAATAAAACAATAAATTCCAGGATTGG 3596
              ||    ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        218 CAGTGAGCAAAGCCAACCTCCTCAATTTCATTTAATAAAACAATAAATTCCAGGATTGG 159

Qy       3597 AAATACAGATACATTACAAAAAGACAGAGCTATCAATATACTGTGTGACCTTATTAC 3656
```

```
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    158 AAATACAGTATACATTACAAAAAGGACAGAGGTCATCAATATACTGTGTGACCTTATTAC 99

Qy   3657 CCCCACTGAGCCCACATATACATGGACCAAGGATGGAACCTTGTTACAGCCCTCAGTAAA 3716
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     98 CCCCAGTGAGGCCACATATACATGGACCAAGGATGGAACCTTGTTACAGCCCTCAGTAAA 39


RESULT 7
BB633755
LOCUS       BB633755                 622 bp    mRNA    linear   EST 26-OCT-2001
DEFINITION  BB633755 RIKEN full-length enriched, adult male spinal cord Mus
            musculus cDNA clone A330077F03 5', mRNA sequence.
ACCESSION   BB633755
VERSION     BB633755.1  GI:16470129
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 622)
  AUTHORS   Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.
            , Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda
            ,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M.,
            Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki
            ,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,
            Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T.,
            Muramatsu,M. and Hayashizaki,Y.
  TITLE     RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/
            Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
            ,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
             Normalization and subtraction of cap-trapper selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res. . 10 (10), 1617-1630 (2000)
             wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
            Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura
            ,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
            Hayashizaki,Y.
             RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer. Genome Res. .
            10 (11), 1757-1771 (2000)
             Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara
            ,Y. and Hayashizaki,Y.
             Computer-based methods for the mouse full-length cDNA
            encyclopedia: real-time sequence clustering for construction of a
            nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
             Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I., Aizawa
            ,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and
            Hayashizaki,Y.
             Computational Analysis of Full Length Mouse cDNAs Compared with
            Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
             Please visit our web site (http://genome.gsc.riken.go.jp) for
            further details.
             e mouse tissues.
FEATURES             Location/Qualifiers
     source          1..622
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="A330077F03"
                     /clone_lib="RIKEN full-length enriched, adult male spinal
                     cord"
                     /sex="male"
                     /tissue_type="spinal cord"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Site_1: SalI; Site_2: BamHI; cDNA library was
                     prepared and sequenced in Mouse Genome Encyclopedia
                     Project of Genome Exploration Research Group in Riken
                     Genomic Sciences Center and Genome Science Laboratory in
                     RIKEN. Division of Experimental Animal Research in Riken
                     contributed to prepare mouse tissues. 1st strand cDNA was
                     primed with a primer [5'
                     GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTTVN 3'], cDNA was
                     prepared by using trehalose thermo-activated reverse
                     transcriptase and subsequently enriched for full-length by
                     cap-trapper. cDNA went through one round of normalization
                     to Rot = 10.0 and subtraction to Rot = 459.0. Second
                     strand cDNA was prepared with the primer adapter of
                     sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCC
                     3']. cDNA was cleaved with XhoI and BamHI. Vector: a
                     modified pBluescript KS(+) after bulk excision from Lambda
                     FLC I. "
BASE COUNT      151 a    157 c    172 g    142 t
ORIGIN

  Query Match              5.8%,  Score 293.6,  DB 10,  Length 622;
  Best Local Similarity   69.1%,  Pred. No. 1.7e-70;
  Matches  401,  Conservative    0;  Mismatches  174;  Indels    5;  Gaps    0;

Qy    186 CCACACCTCAACAAACACTCGTTCACATCAAGACAAAGATCCCAACTTCCATGCTTCCCC 245
          ||  || || || |   | || ||||| ||||||   || |||  ||||||||  |||||
Db     25 CCTGAGTTCCAGGACTGCACGGCTCAGAGGAAGACCGTGAAGGCCTCTGGGATGCCTGGGG 84

Qy    246 CGACTGGAGTGACTGCTCCCGGACCTGTGGGGGAGGAGCATCATATTCTCTGCGGAGATG 305
          |    ||||| || || ||| | |||||||| || || || || ||||| ||| |   | ||
Db     85 CCCTTGGAGCGAGTGTTCCAGAACCTGTGGTGGGGGTGCCTCCTATTCCCTGAGACGCTG 144

Qy    306 TTTGACTGGAAGGAATTGTGAAGGGCAGAACATTCGGTACAAGACATGCAGCAATCATGA 365
            |||    | | ||  |||||||||    || ||||| || |   |||||||| |||    ||
Db    145 CCTGAGCAGCAAGAGCTGTGAAGGGAGAAATATTCGATATAGAACATGCAGTAATGTGGA 204

Qy    366 CTGCCCTCCAGATGCAGAAGATTTCAGACCCAGCACTCCTTCAGCCTACAATGATGTCCA 425
          |||||| ||||| ||||  |||||| |||| || ||||| || ||    | || || ||| |
Db    205 CTGCCCACCAGAAGCAGGTGATTTCCCAGCTCAACACTCTTCTCCTCATAACGACCTCAA 264

Qy    426 GTATCAGGGGCATTACTATGAATGGCTTCCACGATATAATGATCCTGCTGCCCCGTGTGC 485
          ||| || || ||     ||||||||||||||     || |||||| || |        || ||   |
Db    265 GTACCATGGACAGCTTTATGAATGGCTTCTTGTATCTAATGACCCCCACAATCCATCCTC 324

Qy    486 ACTCAAGTGTCATGCACAAGGACAAAACTTGGTGGTGGAGCTGGCACCTAAGGTACTGGA 545
          ||||||||| || ||| |||||     | | |||| ||||| || ||||| |||||  | ||
Db    325 ACTCAAGTGCCAAGCAAAAGGAACCAGCCTGGTTGTGGAACTAGCACCGAAGGTCTTAGA 384

Qy    546 TGGAACTCGTTGCAACACGGACTCCTTGGACATGTGTATCAGTGGCATCTGTCAGGCAGT 605
          ||| || || ||| | || || ||   |||| ||||| ||||| ||| | || ||     ||
Db    385 TGGTACCCGATGCTATACAGAATCACTGGATATGTGCATCAGCGGCCTATGCCAAATTGT 444

Qy    606 GGGCTGCGATCGGCAACTGGGAAGCAATGCCAAGGAGGACAACTGTGGAGTCTGTGCCGG 665
           ||||| ||||  || ||||| |||| || ||| ||||| ||||||||||| ||    |||
Db    445 TGGCTGTGATCACCACCTCCGGAGCACTCTCAAAGAGGATAACTGTGGAGTGTGCCACGG 504

Qy    666 CGATGGCTCCACCTGCAGGCTTGTACGGGGACAATCAAAGTCACACGTTTCTCCTGAAAA 725
           |||||||| || ||| |||| || || || || |   || || ||   | ||| |     ||
Db    505 AGATGGCTCAACGTGCCGGCTGGTCCGAGGGCAGTATAAATCTCAGCTCTCTGCAAGTAA 564

Qy    726 AAGAGAAGAAAATGTAATTGCTGTTCCTTTGGGAAGTCGA 765
          |   || || |  ||  | ||  ||||||  || ||  ||
Db    565 ATCGGATGACACCGTGGTCGCCATTCCTTACGGCAGCAGA 604


RESULT 8
AQ743958
LOCUS       AQ743958                 890 bp    DNA     linear   GSS 16-JUL-1999
DEFINITION  HS_5504_A1_D01_T7A RPCI 11 Human Male BAC Library Homo sapiens
            genomic clone Plate 1080 Col 1 Row G, DNA sequence.
ACCESSION   AQ743958
```

```
VERSION     AQ743958.1  GI:5521486
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 890)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
  MEDLINE   99380589
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            407 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/order_info_tab.htm)
            or from Research Genetics (info@resgen.com). BAC end Web Server:
            http://www.htsc.washington.edu
            Plate: 1080  row: G  column: 1
            Seq primer: T7
            Class: BAC ends
            High quality sequence stop: 890.
FEATURES             Location/Qualifiers
     source          1..890
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="Plate=1080 Col=1 Row=G"
                     /clone_lib="RPCI-11 Human Male BAC Library"
                     /sex="male"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     Male blood DNA was isolated from one randomly chosen donor
                     and partially digested with a combination of EcoRI and
                     EcoRI Methylase. Size selected DNA was cloned into the
                     pBACe3.6 vector at EcoRI sites"
BASE COUNT      228 a    185 c    218 g    259 t
ORIGIN

  Query Match             5.6%;  Score 281.8;  DB 17;  Length 890;
  Best Local Similarity   97.6%;  Pred. No. 4.2e-67;
  Matches  286;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       1696 CCAACTTTATTCAGAACCTGGTCAGCTGCAGTACCACGTGTGGGCAGGTGTGCAG 1755
              || | ||||||||||||||||  |||||||||||||||||||||||||| ||||||
Db        125 CCTAGGTTCATTCAGAACCTGGTCAGCTGCAGTACTACGTGTGGGCTGGTGTGCAG 184

Qy       1756 GTCCGCGAGGTGAAGTGCCGCGTGCTCCTCACATTCACGCAGACTGAGACGAGCTGCCC 1815
              ||||| || ||||||||||||||||||||||||||||||||||| | | ||||||||
Db        185 GTCCGTGAGGTGAAGTGCCGCGTGCTCCTCACATTCACGCAGACTGATATTGAGCTGCCC 244

Qy       1816 GAGGAAGAGTGTGAAGGCCCCAAGTGCCCACCGAACGGCCCTGCCTCCTGGAAGCATGT 1875
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        245 GAGGAAGATTGTGAAGGCCCCAAGTGCCCACCGAACGGCCCTGCCTCCTGGAAGCATGT 304

Qy       1876 GATGAGAGCCGGCCTCCCGGAGAGCTAGACATCCCTCTCCCTGAGGACAGTGAGACGACT 1935
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        305 GATGAGAGCCGGCCTCCCGGAGAGCTAGACATCCCTCTCCCTGAGGACAGTGAGACGACT 364

Qy       1936 TACGACTGGGAGTACGCTGGGTTCACCCCTGCACAGCAACATGCTGGGAGG 1988
              ||||||||| |||||||||||||||||| |||||||||||| |||||||
Db        365 TACGACTGGGAGTACGCTGGGTTCACCCCTGCACAGCAACATGCTGGGAGG 417


RESULT 9
BE888902
LOCUS       BE888902                 850 bp    mRNA    linear   EST 20-OCT-2000
DEFINITION  601513828F1 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:3915294 5',
            mRNA sequence.
ACCESSION   BE888902
VERSION     BE888902.1  GI:10345670
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 850)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM9738  row: g  column: 07
            High quality sequence stop: 742.
FEATURES             Location/Qualifiers
     source          1..850
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3915294"
                     /clone_lib="NIH_MGC_71"
                     /tissue_type="leiomyosarcoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
                     Average insert size 2.1 kb. "
BASE COUNT      202 a    219 c    228 g    201 t
ORIGIN

  Query Match             5.5%;  Score 278;  DB 12;  Length 850;
  Best Local Similarity   67.3%;  Pred. No. 4.7e-66;
  Matches  451;  Conservative    0;  Mismatches  215;  Indels    4;  Gaps    4;

Qy        186 CCAGACTTCAAGAAACACTCGTTCAGATGAAGACAAAGATGGCAACTGGGATGCTTGGGG 245
              |  ||    || ||  |  |  |  ||  ||  ||  |||    || ||     ||||||| ||||
Db        103 CCTGAGTTCCAGGACCGCACGCTCTGAGGAGGACGGGACGGCCTATGGGATGCCTGGGG 162

Qy        246 CGACTGGACTGACTGCTCCCGGACCTGTGGGGGAGGAGCATCATATCCTCTCGGAGATG 305
              |   |||||||| |||| || ||||| ||||| || || || || ||||| || | ||
Db        163 CCCATGGACTGAATGCTCAGCACCTGCGGGGTGGGCCTCCTACTCTCTGAGGCGCTG 222

Qy        306 TTCGACTGGAAGGAATGTGAAGGCAGAACATTCGGTACAAGACATCCAGCAATCATGA 365
              |||    | | ||  |||||||     || || || |||  ||||||| |||    ||
Db        223 CCTGACCAGCAAGAGCTGTGAAGGAAGAAATATCCGATACAGAACATGCAGTAATCTGGA 282

Qy        366 CTGCCCTCCAGATGGAGAAGATTCAGAGCCCAGCAGTCCTTACCTTACAACGATGTCCA 425
              ||  ||  |||||  ||||  ||||| |||| ||||| |||||||   |   |||||||| |
Db        283 CTGCCCACCAGAAGCAGGTGATTTCGAGCTCAGCAATGCTCAGCTCATAATGATGTTAA 342

Qy        426 GTATCAGGGCATTACTATGAATGGCTTCCACGATATAATGATCCTGCTGCCCCGTGTGGC 485
              | | || || || |  ||||||||||||||      | ||||| ||||      ||| ||| |
Db        343 GCACCATGGCCAGTTTATGAATGGCTTCCTGTGTCTAATGACCCTGACAACCCATGTTTC 402

Qy        486 ACTCAAGTGTCATGCACAAGACAAAACTTGGTGGTGGAGCAGGCACCTAAGGTACTGGA 545
              ||||||||| ||  |  |||||  || | |||| || || || ||||||||||  | ||
Db        403 ACTTAAGTGCCAAGCCAAAGGAACAACCCTGGGTGTTGAACTACCACCTAAGGTCTTAGA 462

Qy        546 TGGAACTCGTTGCAACACGGACTCCTTGGACATGTGTATCACGGCACTCGTCAGTCAGT 605
              ||| || || ||| | || || || ||||| |||| |||||||  | || ||     ||
Db        463 TGGTACGCGCTGCTATACAGAATCTTTGGATATGTGCATCAGTGGTTTATGCAAATTGT 522

Qy        606 GGGCTGCGATCGGCAACTGGGAAGCAATGCCAAGGAGGACAACCGTGGAGTCTGTGGCGG 665
               ||||||||||   ||  |||||||||||   |  ||||| || ||||||||  ||||    ||
Db      522 TGGCTGCGATCACCAGCTCCGAAGCACCGTCAAGGAAGATAACTGTCCCCTCTCCAACCC 581

Qy      666 CGAT-GGCTCCACCTGCAGGCTTGTACGGGGACAATCAAAGTCACACGTTTC-TCCTGAA 723
             ||| ||  ||||||||| |||| || || || ||| |  || || ||  | ||  |
Db      582 AGATGGGGTCCACCTGCCGGCTGGTCCGACCCACTATAAATCCCAGCTCTCAGCAACCC 641

Qy      724 AAAAGAGAAGAAAATGTAATTGC-TGTTCCTTTGGGAAGTCGAAGTGTGAGAATTACAGT 782
            ||   | || || | |||  ||||     |||| | || |     | | |     ||    |
Db      642 AATCGGGATGATACTGTGGTTGCAAATTCCCTATGGGAAGTTGACTATTTCGCTTGGTTT 701

Qy      783 GAAAGGACCTGCCCACCTCTTTATTGAATCAAAAACACTTCAAGGCAAGCAAAGGAGAACA 842
             ||||| ||||  ||| | ||  | | | | |||||         |||  ||| |  | |
Db      702 AAAAGGGCCTGATCACTTTTTATTGGGAACCAAAACCTCCACGGCACTCAAGGTGGCAAC 761

Qy      843 CAGCTTTAAC 852
              | | || |
Db      762 AGGGTCTATC 771


RESULT 10
BB478282
LOCUS       BB478282                619 bp    mRNA    linear   EST 25-OCT-2001
DEFINITION  BB478282 RIKEN full-length enriched, 13 days embryo heart Mus
            musculus cDNA clone D330029I02 3', mRNA sequence.
ACCESSION   BB478282
VERSION     BB478282.2  GI:16440420
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 619)
  AUTHORS   Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A
            ., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda
            ,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M.,
            Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki
            ,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,
            Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T.,
            Muramatsu,M. and Hayashizaki,Y.
  TITLE     RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
  JOURNAL   Unpublished (2001)
COMMENT     On Jul 23, 2000 this sequence version replaced gi:9395891.
            Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/
            Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
            ,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
             Normalization and subtraction of cap trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res. . 10 (10), 1617-1630 (2000)
             Sasaki,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
            Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura
            ,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
            Hayashizaki,Y.
             RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer. Genome Res. .
            10 (11), 1757-1771 (2000)
             Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara
            ,Y. and Hayashizaki,Y.
             Computer-based methods for the mouse full-length cDNA
            encyclopedia: real-time sequence clustering for construction of a
            nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
             Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I., Aizawa
            ,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and
            Hayashizaki,Y.
             Computational Analysis of Full-Length Mouse cDNAs Compared with
            Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
             Please visit our web site (http://genome.gsc.riken.go.jp/) for
            further details.
            cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues.
FEATURES             Location/Qualifiers
     source          1..619
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="D330029I02"
                     /clone_lib="RIKEN full-length enriched, 13 days embryo
                     heart"
                     /tissue_type="heart"
                     /dev_stage="13 days embryo"
                     /lab_host="DH10B"
                     /note="Site_1: SalI; Site_2: BamHI; cDNA library was
                     prepared and sequenced in Mouse Genome Encyclopedia
                     Project of Genome Exploration Research Group in Riken
                     Genomic Sciences Center and Genome Science Laboratory in
                     RIKEN. Division of Experimental Animal Research in Riken
                     contributed to prepare mouse tissues. 1st strand cDNA was
                     primed with a primer [5'
                     GAGAGAGAGAGCGGCCGCAACTCGAGTTTTTTTTTTTTTTTTTTVN 3'], cDNA was
                     prepared by using trehalose thermo-activated reverse
                     transcriptase and subsequently enriched for full-length by
                     cap-trapper. Second strand cDNA was prepared with the
                     primer adapter of sequence [5'
                     GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCC 3']. cDNA
                     was cleaved with BamHI and XhoI. Vector: a modified
                     pBluescript KS(+) after bulk excision from Lambda FLC I."
BASE COUNT      160 a    142 c    169 g    148 t
ORIGIN

  Query Match              5.4%;  Score 274;  DB 10;  Length 619;
  Best Local Similarity   70.8%;  Pred. No. 5.3e-65;
  Matches  364;  Conservative    0;  Mismatches  150;  Indels    0;  Gaps    0;

Qy      215 AAGACAAAGATGGCAACTGGGATGCTTGGGGCGACTGGAGTGACTGCTCCGGACCTGTG 274
            |||||    || |||  ||||||||||||||||    ||||| || || ||| | ||||||
Db        1 AAGACCGTGAAGGCTTCTGGGATGCTTGGGGCCCTTGGAGCGAGTGTTCCAGAACCTGTG 60

Qy      275 CGGCAGGAGCATCATATTCTCTGCGGAGATGTTTGACTGGAAGGAATTGTGAAGGGCAGA 334
            | || || || || ||||| ||| |   | || ||||    | | ||  ||||||||    |
Db       61 GTGGGGGTGCCTCCTATTCCCTGAGACGCTGCTTGAGCAGCAAGAGCTGTGAACCCACAA 120

Qy      335 ACATTCGGTACAAGACATGCAGCAATCATGACTGCCCTCCAGATGCAGAAGATTTCAGAG 394
            | ||||| || |  |||||||| |||     ||||||| ||||| ||||    |||||| |||
Db      121 ATATTCGATATAGAACATGCAGTAATGTGGACTGCCCACCAGAAGCAGGTGATTTCCGAG 180

Qy      395 CCCAGCAGTGCTCAGCCTACAATGATGTCCAGTATCAGGGGCATTACTATGAATGGCTTC 454
            | || ||||| || ||   | || || ||| |||| || || ||      ||||||||||||
Db      181 CTCAACAGTGTTCTGCCTCATAACGACGTCAAGTACCATGGACAGCTTTATGAATGCCTTC 240

Qy      455 CACGATATAATGATCCTGCTGCCCCGTGTGCACTCAAGTGTCATGCACAAGGACAAAACT 514
            |    || ||||||| || |        || ||  |||||||||| || ||| |||||    | |
Db      241 CTGTATCTAATGACCCCGACAATCCATGCTCACTCAAGTGCCAAGCAAAAGGAACCAGCC 300

Qy      515 TGGTGGTGGAGCTGGCACCTAAGGTACTGGATGGAACTCGTTGCAACACGGACTCCTTGG 574
            |||| ||||| || ||||| |||||    | ||||| || || ||| | || || ||   |||
Db      301 TGGTTGTGGAACTAGCACCGAAGGTCTTAGATGGTACGCCGATGCTATACAGAATCACTGG 360

Qy      575 ACATGTGTATCAGTGGCATCTGTCAGGCAGTGGGCTGCGATCGGCAACTGGGAAGCAATG 634
            | ||||| ||||| ||| | || ||       || ||||| ||||   || ||||| |||| ||
Db      361 ATATGTGCATCAGCGGCCTATGCCAAATTGTTGGCTCTGATCACCAGCTCCGGACCACTC 420

Qy      635 CCAAGGAGGACAACTGTGGAGTCTGTGCCGGCGATGGCTCCACCTGCAGGCTTGTACGGG 694
             ||| ||||| |||||||||| ||     ||| |||||||| || ||| |||| || || |
Db      421 TCAAAGAGGATAACTGTGGAGTGTGCAACGGAGATGGCTCAACGTGCCGGCTGGTCCGAG 480

Qy      695 GACAATCAAAGTCACACGTTTCTCCTGAAAAAAG 728
            | || |   || || ||  | ||| |     ||| |
Db        481 GGCAGTATAAATCTCAGCTCTCTGCAAGTAAATG 514


RESULT 11
BE162337/c
LOCUS       BE162337                 271 bp    mRNA    linear   EST 21-JUN-2000
DEFINITION  PM2-HT0451-060100-003-b04 HT0451 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BE162337
VERSION     BE162337.1  GI:8625058
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 271)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=cd&t2=PM2-HT0451-060
            100-003-b04&t3=2000-01-08&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 11
            High quality sequence stop: 271.
FEATURES             Location/Qualifiers
     source          1..271
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="HT0451"
                     /dev_stage="Adult"
                     /note="Organ: head_neck; Vector: pUC18; Site_1: SmaI;
                     Site_2: SmaI; A mini library was made by cloning products
                     derived from ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 - Ludwig Institute for Cancer Research)
                     profiles into the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
BASE COUNT       63 a     63 c     58 g     87 t
ORIGIN

  Query Match              5.3%;  Score 267.8;  DB 10;  Length 271;
  Best Local Similarity   99.3%;  Pred. No. 1.9e-63;
  Matches  269;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy        730 GAAGAAAATGTAATTGCTGTTCCTTTGGGAAGTCGAAGTGTGAGAATTACAGTGAAAGGA 789
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        271 GAAGAAAATGTAATTGCTGTTCCTTTGGGAAGTCGAAGTGTGAGAATTACAGTGAAAGGA 212

Qy        790 CCTGCCCACCTCTTTATTGAATCAAAAACACTTCAAGGAAGCAAAGGAGAACACAGCTTT 849
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        211 CCTGCCCACCTCTTTATTGAATCAAAAACACTTCAAGGAAGCAAAGGAGAACACAGCTTT 152

Qy        850 AACAGCCCCGGCGTCTTTTGTCTGAGAAAACACAACAGTGGAATTTCAGAGGGCTCCGAG 909
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        151 AACAGCCCCGGCGTCTTTTGTCTGAGAAAACACAACAGTGGAATTTCAGAGGGCTCCGAG 92

Qy        910 AGGCAAACTTTTAAGATTCCAGGACCTCTGATGGCTGATTTCATCTTCAAGACCAGGTAC 969
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         91 AGGCAAACTTTTAAGATTCCAGGACCTCTGATGGCTGATTTCATCTTCAAGACCAGGTAC 32

Qy        970 ACTGCAGCCAAAGACAGCGTGAATCAGTTCT 1000
              |||||||||||||||||||||||||||||||
Db         31 ACTGCAGCCAAAGACAGCGTGAATCAGTTCT 1


RESULT 12
AK020115
LOCUS       AK020115                 979 bp    mRNA    linear   HTC 19-JAN-2002
DEFINITION  Mus musculus 12 days embryo male wolffian duct includes surrounding
            region cDNA, RIKEN full-length enriched library,
            clone:6720426B09:thrombospondin type 1 domain containing protein,
            full insert sequence.
ACCESSION   AK020115
VERSION     AK020115.1  GI:12860599
KEYWORDS    HTC; CAP trapper.
SOURCE      Mus musculus (strain:C57BL/6J) 12 days embryo male wolffian duct
            includes surrounding region cDNA to mRNA, clone_lib:RIKEN
            full-length enriched mouse cDNA library clone:6720426B09.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Carninci,P. and Hayashizaki,Y.
  TITLE     High-efficiency full-length cDNA cloning
  JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
  MEDLINE   99279253
   PUBMED   10349636
REFERENCE   2
  AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
  TITLE     Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
  JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
  MEDLINE   20499374
   PUBMED   11042159
REFERENCE   3
  AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
            Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
            Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
            Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
            Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
            Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
            Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
  TITLE     RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer
  JOURNAL   Genome Res. 10 (11), 1757-1771 (2000)
  MEDLINE   20530913
   PUBMED   11076861
REFERENCE   4
  AUTHORS   Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y.,
            Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S.,
            Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I.,
            Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R.,
            Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T.,
            Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H.,
            Kuehl,P., Lewis,S., Matsuo,Y., Nikaido,I., Pesole,G.,
            Quackenbush,J., Schriml,L.M., Staubli,F., Suzuki,R., Tomita,M.,
            Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H.,
            Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N.,
            Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C.,
            Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D.,
            Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P.,
            Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P.,
            Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H.,
            Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H.,
            Toyo-oka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L.,
            Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohtsuki,S.
            and Hayashizaki,Y.
  TITLE     Functional annotation of a full-length mouse cDNA collection
  JOURNAL   Nature 409 (6821), 685-690 (2001)
  MEDLINE   21085660
   PUBMED   11217851
REFERENCE   5  (bases 1 to 979)
  AUTHORS   Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A.,
            Arakawa,T., Baldarelli,R., Bono,H., Brownstein,M , Bult,C ,
            Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T.,
            Hara,A., Hayatsu,N., Hill,D., Hiramoto,K., Hiraoka,T., Hori,F ,
            Hume,D., Imotani,K., Ishii,Y , Itoh,M , Izawa,M., Kasukawa,T.,
            Kato,H , Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S.,
            Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K.,
            Numazaki,R., Ohno,M., Okazaki,Y , Okido,T., Owa,C., Quackenbush,J ,
            Saito,H., Saito,R., Sakai,C , Sakai,K , Sano,H , Sasaki,D.,
            Schriml,L., Shibata,K , Shibata,Y., Shinagawa,A., Shiraki,T.,
            Sogabe,Y., Suzuki,H , Tagami,M., Tagawa,A., Takahashi,F.,
            Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yamanaka,I.,
            Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and
            Hayashizaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (18 AUG 2000) Yoshihide Hayashizaki, The Institute of
            Physical and Chemical Research (RIKEN), Laboratory for Genome
            Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
            RIKEN Yokohama Institute, 1 7 22 Suehiro cho, Tsurumi-ku, Yokohama,
            Kanagawa 230-0045, Japan (E mail.genome-res@gsc.riken.go.jp,
            URL http://genome gsc riken go jp/, Tel 81 45-503-9222,
            Fax:81-45-503-9216)
COMMENT     Please visit our web site (http://genome gsc riken go.jp/) for
            further details.
            cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues. First strand cDNA was primed with a primer
            [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTTTVN 3'], cDNA was
            prepared by using trehalose thermo-activated reverse transcriptase
            and subsequently enriched for full-length by cap-trapper. Second
            strand cDNA was prepared with the primer adapter of sequence [5'
            GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCC 3']. cDNA was cleaved
            with BamHI and XhoI. Vector: a modified pBluescript KS(+) after
            bulk excision from Lambda FLC I  Cloning sites, 5' end: SalI; 3'
            end: BamHI. Host: DH10B.
FEATURES             Location/Qualifiers
     source          1..979
                     /organism="Mus musculus"
                     /strain "C57BL/6J"
                     /db_xref="FANTOM_DB:6720426B09"
                     /db_xref="MGD:MGI:1911603"
                     /db_xref="taxon:10090"
                     /clone="6720426B09"
                     /sex "male"
                     /tissue_type "wolffian duct includes surrounding region"
                     /clone lib "RIKEN full-length enriched mouse cDNA library"
                     /dev_stage "12 days embryo"
     CDS             85..663
                     /note="Thrombospondin type 1 domain containing protein
                     data source:Pfam, source key:PF00090, evidence:ISS
                     putative"
                     /codon_start=1
                     /protein_id="BAB32000.1"
                     /db_xref "GI:12860600"
                     /db_xref "MGD:MGI:1924989"
                     /translation "MECCRFAAPCTILIVLAFLLLSSFTARSEEEFEGLWLAWGIWSE
                     CSRTCGGGASYSLRRCLSSKSCEGRNIRYRTCSNVDCPPEAGDFRAQQCSAHNDVKYH
                     GQLYFWLPVSNDPDNPCSLKCQAKGISLVVELAPKVLYGTRCYTESLDMCISGLCQIV
                     GCDHQLGSTVKKDNCGVCNGDGSTCRVGPRAV"
BASE COUNT      252 a    226 c    265 g    236 t
ORIGIN

  Query Match             5.3%;  Score 266.8;  DB 11;  Length 979;
  Best Local Similarity   70.2%;  Pred. No. 7e-63;
  Matches  358;  Conservative    0;  Mismatches  152;  Indels    0;  Gaps    0;

Qy        186 CCAGACCTCAAGAAACACTCCTTCAGATGAAGACAAACATGGCAACTCCCATGCTTGGGC 245
              || ||  || || |   | || ||||| ||||||   || |||  ||||||||| |||||
Db        144 CCTGAGTTCCAGGACTGCACGCTCAGAGGAAGACCGTGAAGCCCTCTGGCATGCCTGGGG 203

Qy        246 CGACTGGAGTGACTGCTCCCGGACCTGTGGGGGAGGAGCATCATATTCTCTGCGGAGATG 305
              |   ||||| || || ||| | |||||||| || || || || ||||| ||| |  | ||
Db        204 CCCTTGGAGCGAGTGTTCCAGAACCTGTGGTGGGGTGCCTCCTATTCCCTGAGACGCTG 263

Qy        306 TTTGACTGGAAGGAATTGTGAAGGGCAGAACATTCGGTACAAGACATGCAGCAATCATGA 365
                |||   | | ||  |||||||||   || ||||| || |  |||||||| |||   ||
Db        264 CCTGAGCAGCAAGAGCTGTGAAGGGAGAAATATTCGATATAGAACATGCAGTAATGTGGA 323

Qy        366 CTGGCTCCACATGCAGAAGATTTCAGAGCCCAGCACTGCTCAGCCTACAATGATGTCCA 425
              |||||| ||||| ||||  |||||| |||| || ||||| || ||   | || || ||| |
Db        324 CTGCCCACCAGAAGCAGGTGATTTCCGAGCTCAACAGTGTTCTGCTCATAACGACGTCAA 383

Qy        426 GTATCAGGGGCATTACTATGAATGGCTTCCACGATATAATGATCCTGCTGCCCCGTGTGC 485
              ||| || || ||    ||||||||||||||   || |||||| || |     || ||  |
Db        384 GTACCATGGACAGCTTTATGAATGGCTTCCTGTATCTAATGACCCCGACAATCCATGCTC 443

Qy        486 ACTCAAGTGTCATGCACAAGGACAAAACTTGGTGGTGGAGCTGGCACCTAAGGTACTGGA 545
              ||||||||| || ||| |||||   | | |||| ||||| || ||||| |||||  |  |
Db        444 ACTCAAGTGCCAAGCAAAAGGAACCAGCTTGGTTGTGGAACTAGCACCGAAGGTCTTATA 503

Qy        546 TGGAACTCGTTGCAACACGGACTCCTTGGACATGTGTATCAGTGGCATCTGTCAGGCAGT 605
              ||| || || ||| | || || ||  |||| ||||| ||||| ||| | || ||    ||
Db        504 TGGTACGCGATGCTATACAGAATCACTGGATATGTGCATCAGCGGCCTATGCCAAATTGT 563

Qy        606 GGGCTGCGATCGGCAACTGGAAGCAATGCCAAGGAGGACAACTGTGGAGTCTGTGCCGG 665
               ||||| ||||  || ||||| |||| || |||  |||| ||||||||||| ||   |||
Db        564 TGGCTGTCATCACCAGCTGCCCAGCACTGTCAAAAAGGATAACTGTGGACTGTGCAACGG 623

Qy        666 CGATGGCTCCACCTGCAGGCTTGTACGGGG 695
               |||||||| || ||| || |||  | | |
Db        624 AGATGGCTCAACGTGCCGGGTTGGTCCGAG 653


RESULT 13
BB659292
LOCUS       BB659292                 660 bp    mRNA    linear   EST 26-OCT-2001
DEFINITION  BB659292 RIKEN full-length enriched, 13 days embryo heart Mus
            musculus cDNA clone D330029I02 5', mRNA sequence.
ACCESSION   BB659292
VERSION     BB659292.1  GI:16493113
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 660)
  AUTHORS   Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.
            , Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda
            ,M., Koya,S , Matsuyama,T , Miyazaki,A., Nomura,K., Ohno,M.,
            Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki
            ,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,
            Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T.,
            Muramatsu,M. and Hayashizaki,Y.
  TITLE     RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
  JOURNAL   Unpublished (2001)
COMMENT     Contact. Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi ku, Yokohama, Kanagawa 230 0045, Japan
            Tel: 81 45 503 9222
            Fax. 81 45 503-9216
            Email. genome res@gsc.riken.go.jp,
            URL http://genome gsc riken go jp/
            Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
            ,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
             Normalization and subtraction of cap trapper selected cDNAs to
            prepare full length cDNA libraries for rapid discovery of new
            genes. Genome Res. , 10 (10), 1617 1630 (2000)
             wagi,K., Fujiwake,S., Inoue,K., Iotawa,Y., Izawa,M., Ohara,E.,
            Watahiki,M , Yoneda,Y , Ishikawa,T , Ozawa,K , Tanaka,T , Matsuura
            ,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
            Hayashizaki,Y.
             RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer. Genome Res.
            10 (11), 1757-1771 (2000)
             Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara
            ,Y. and Hayashizaki,Y.
             Computer-based methods for the mouse full-length cDNA
            encyclopedia: real-time sequence clustering for construction of a
            nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
             Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I., Aizawa
            ,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and
            Hayashizaki,Y.
             Computational Analysis of Full-Length Mouse cDNAs Compared with
            Human Genome Sequences. Mamm. Genome 12, 673-677 (2001)
             Please visit our web site (http://genome.gsc.riken.go.jp/) for
            further details.
             e mouse tissues.
FEATURES             Location/Qualifiers
     source          1..660
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="1330029I02"
                     /clone_lib="RIKEN full-length enriched, 13 days embryo
                     heart"
                     /tissue_type="heart"
                     /dev_stage="13 days embryo"
                     /lab_host="DH10B"
                     /note="Site_1: SalI; Site_2: BamHI; cDNA library was
                     prepared and sequenced in Mouse Genome Encyclopedia
                     Project of Genome Exploration Research Group in Riken
                     Genomic Sciences Center and Genome Science Laboratory in
                     RIKEN. Division of Experimental Animal Research in Riken
                     contributed to prepare mouse tissues. 1st strand cDNA was
                     primed with a primer [5'-
                     GAGAGAGAGAGAGAGCGCAACTCGAGTTTTTTTTTTTTTTTTVN-3']. cDNA was
                     prepared by using trehalose thermo-activated reverse
                     transcriptase and subsequently enriched for full-length by
                     cap-trapper. Second strand cDNA was prepared with the
                     primer adapter of sequence [5'-
                     GAGAGAGAGATCTCGAGTTAATTAAATTAATCCCCCCCCCCCC-3']. cDNA
                     was cleaved with BamHI and XhoI. Vector: a modified
                     pBluescript KS(+) after bulk excision from Lambda FLC I."
BASE COUNT      161 a    163 c    191 g    145 t
ORIGIN

  Query Match              5.2%;  Score 262.4;  DB 10;  Length 660;
  Best Local Similarity   70.6%;  Pred. No. 9.8e-62;
  Matches  350;  Conservative    0;  Mismatches  146;  Indels    0;  Gaps    0;

Qy        186 CCAGACCTCAAGAAACACTGGTTCAGATGAAGACAAAGATGGCAACTGGGATGCTTGGGT 245
              || ||  || || |   | || ||||| ||||||    || |||   |||||||| |||||
Db        164 CCTGAGTTCCAGGACTGCAGGCTAGAGGAAGACCGTGAAGGCCTCTGGGATGCCTGGGG 223

Qy        246 CGACTGGACTGACTGCTCCCGGACCTGTGGGGGAGGAGCATCATATTCTCTGCGGAGATG 305
              |   ||||| || || ||| | ||||||| || || || || ||||| ||  |  | ||
Db        224 CCCTTGGAGCGAGTGTTCCAGAACCTGTGGTGGGGGTGCCTCCTATTCCCTGAGACGCTG 283

Qy        306 TTTGACTGGAAGGAATTGTGAAGGGCAGAACATTCCGTACAAGACATGCAGCAATCATGA 365
                |||   | | ||  |||||||||   || ||||| || |  ||||||| |||    |
Db        284 CCTGAGCAGCAAGAGCTGTGAAGGGAGAAATATTCGATATAGAACATGCAGTAATCTTGA 343

Qy        366 CTGCCCTCCAGATGCAGAAGATTTCAGAGCCAGCAGTGCTCAGCCTACAATGATGTCCA 425
              |||||| ||||| ||||  |||||| |||| || ||||| || ||   | || || ||| |
Db        344 CTGCCCACCAGAAGCAGGTGATTTCGAGCTCAACAGTGTTCTGCTCATAACGACGTCAA 403

Qy        426 GTATCAGGGCATTACTATGAATGGCTTCCACGATATAATGATCCTGCTGCCCCGTGTGC 485
              ||| || | ||     |||||||||||||   || |||||| || |       | || |
Db        404 GTACCATGGACAGCTTTATGAATGGCTTCCTGTATCTAATGACCCGACAATCCATGCTC 463

Qy        486 ACTCAACGTGTCATGCACAAGGACAAAACTTGGTGGTGGAGCTGGCACCTAAGGTACTGGA 545
              |||||||||| || ||| |||||    | | |||| ||||| || ||||| |||||    | ||
Db        464 ACTCAACGTGCCAAGCAAAAGGAACTAGCCTGGTTGTGGAACTAGCACCGAAGGTCTTAGA 523
```

```
Qy        546 TGGAACTCGTTGCAACACGGACTCCTTGGACATGTGTATCAGTGGGATCTGTCAGGCAGT 605
              ||| ||| || ||| | || |  ||  |||| ||||| ||||| ||| | || ||      ||
Db        524 TGGTACGCGATGCTATACAGAATCACTGGATATGTGCATCAGGGCCTATGCTAAATTGT 583

Qy        606 [illegible]AGTCTGTGCCGG 665
               ||||| |||| |    | ||||| |||| || ||| || || |||||||| | ||   |||
Db        584 TGGCTGTGAC[illegible]AAAGAAGAAGCAA[illegible]AACGG 643

Qy        666 CGATGGCTCCACCTGC 681
               |||||||| || |||
Db        644 AGATGGCTCAACGTGC 659


RESULT 14
BF543337/c
LOCUS       BF543337                 408 bp    mRNA    linear   EST 11-DEC-2000
DEFINITION  UI-R-Y0-abq-a-03-0-UI.r1 UI-R-Y0 Rattus norvegicus cDNA clone
            UI-R-Y0-abq-a-03-0-UI 5', mRNA sequence.
ACCESSION   BF543337
VERSION     BF543337.1  GI:11634444
KEYWORDS    EST.
SOURCE      Norway rat.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 408)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
  MEDLINE   97044477
COMMENT     Contact: Soares, MB
            Program for Rat Gene Discovery and Mapping
            University of Iowa
            451 Eckstein Medical Research Building Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565
            Email: msoares@blue.weeg.uiowa.edu
            cDNA Library Preparation: M.B. Soares Lab Clone distribution:
            clones will be available through Research Genetics (www.resgen.com)
            This clone is also available through the I.M.A.G.E. Consortium at
            LLNL (info@image.llnl.gov). IMAGE ID  1790944
            Seq primer: M13 Forward.
FEATURES             Location/Qualifiers
     source          1..408
                     /organism="Rattus norvegicus"
                     /strain="Sprague-Dawley"
                     /db_xref="taxon:10116"
                     /clone="UI-R-Y0-abq-a-03-0-UI"
                     /clone_lib="UI-R-Y0"
                     /dev_stage="adult"
                     /lab_host="DH10B (Life Technologies)"
                     /note="Vector: pT7T3D Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; The UI-R-Y0
                     library is a subtracted library derived from an
                     individually-tagged normalized whole eye (minus the lens)
                     library. The driver for the subtraction consisted of a
                     pool of all previous libraries (UI-R-A0, UI-R-A1, UI-R-E0,
                     UI-R-E1, UI-R-C0, and UI-R-C1). The tag is a string of
                     3-5 nucleotides present between the Not I site and the
                     oligo-dT track which allows identification of the library
                     of origin of a clone within the mixture. The subtracted
                     library (UI-R-Y0) was constructed as follows: PCR
                     amplified cDNA inserts from previous library clones from
                     which 3' ESTs had been derived were used as a driver in a
                     hybridization with the normalized whole eye library in
                     the form of single-stranded circles. The remaining
                     single-stranded circles (subtracted library) was purified
                     by hydroxyapatite column chromatography, converted to
                     double-stranded circles and electroporated into DH10B
                     bacteria (Life Technologies) to generate the UI-R-Y0
                     library. This procedure has been previously described
                     (Bonaldo, Lennon and Soares, Genome Research 6: 791-806,
                     1996)"
BASE COUNT       91 a    113 c    105 g     99 t
ORIGIN

  Query Match             5.1%;  Score 259.6;  DB 12;  Length 408;
  Best Local Similarity   85.5%;  Pred. No. 4.7e-61;
  Matches  289;  Conservative    0;  Mismatches   49;  Indels    0;  Gaps    0;

Qy          1 ATGGCTTCCTGGACGAGCCCCTGGTGGGTGCTGATACGGCATGGTCTTCATGCACTCTCCC 60
              ||||||| |||||    |   ||||||||| |||||| | | ||| |||| |     ||||
Db        348 ATGGCTTTCTGGAAAGGATCCTGGTGGCTGCTGACACGAATGATCTTTACAGTCTCTGGT 289

Qy         61 CTCCCGCAGACCACAGCTGAGAAATCTCCTGGAGCCTATTTCCTTCCCGAGTTTGCACTT 120
              |||| |||| ||||||| ||||| |||||||||||||| ||||||||| ||||| |||||
Db        288 CTCCTGCAGGCCACAGCAGAGAAGTCTCCTGGAGCCTACTTCCTTCCTGAGTTCGCACTC 229

Qy        121 TCTCCTCAGGGAAGTTTTCTGGAAGACACAACAGGGGAGCAGTTCCTCACTTATCGCTAT 180
              ||||||||||||| ||||| ||||| || ||||||||| ||||||||||| || ||||||
Db        228 TCTCCTCAGGGAAGCTTTCTAGAAGATACCACAGGGGAACAGTTCCTCACATACCGCTAT 169

Qy        181 GATGACCAGACCTCAAGAAACACTCGTTCAGATGAAGACAAAGATGGCAACTGGGATGCT 240
              ||||| |||||||||||||||   |||   || |||||||||||||||||||||||||||
Db        168 GATGATCAGACCTCAAGAAATGCTCACTCCGATGAAGACAAAGATGGCAACTGGGATGCC 109

Qy        241 TGGGGCGACTGGAGTGACTGCTCCCGGACCTGTGGGGCAGGAGCATCATATTCTCTGCGG 300
              ||||| |||||||||||||||||||||||||||||||||||||||| || |||||| ||||
Db        108 TGGGGAGACTGGAGTGACTGCTCCCGGACCTGTGGGGCAGGAGCCTCCTATTCTTTGCGA 49

Qy        301 AGATGTTTGACTGGAAGGAATTGTGAAGGGCAGAACAT 338
              |||||| ||||||||||||| |||||||| ||||||||
Db         48 AGATGTCTGACTGGAAGGAACTGTGAAGGACAGAACAT 11


RESULT 15
BI523618
LOCUS       BI523618               581 bp    mRNA    linear   EST 29-AUG-2001
DEFINITION  603051757F1 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5201199 5',
            mRNA sequence.
ACCESSION   BI523618
VERSION     BI523618.1  GI:15348410
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 581)
  AUTHORS   NIH MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11504  row: b  column: 16
            High quality sequence stop: 577.
FEATURES             Location/Qualifiers
     source          1..581
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5201199"
                     /clone_lib="NIH_MGC_122"
                     /lab_host="DH10B"
                     /note="organ: pooled lung and spleen; Vector: pCMV-SPORT6;
                     Site_1: NotI; Site_2: EcoRV (destroyed); RNA source
                     anonymous pool of 24 week female lung, 16 week female
                     spleen, and 20 22 week male spleens.  Library is oligo-dT
                     primed and directionally cloned (EcoRV site is destroyed
                     upon  cloning).  Average insert size 1.4 kb, insert size
                     range 1-3 kb. Library is normalized and enriched for
                     full-length clones and was  constructed by C. Gruber
                     (Invitrogen).  Research Genetics tracking  code 026. Note:
                     this is a NIH_MGC Library."
BASE COUNT      140 a    151 c    159 g    131 t
ORIGIN

  Query Match             5.0%;  Score 256;  DB 13;  Length 581;
  Best Local Similarity   72.6%;  Pred. No. 5.7e-60;
  Matches  331;  Conservative    0;  Mismatches  125;  Indels    0;  Gaps    0;

Qy        186 CCAGACCTCAAGAAACACTCGTTCAGATGAAGACAAAGATGGCAACTGGGATGCTTGGGG 245
              || ||  || || | | | || || || || |||    || |||   ||||||| |||||
Db        125 CCTGAGTTCCAGGACCGCACGCTCCGAGGAGGACCGGGACGGCCTATGGGATGCCTGGGG 184

Qy        246 CGACTGGAGTGACTGCTCCCGGACCTGTGGGGAGGAGCATCATATTCTCTGCGGAGATG 305
              |    |||||||| ||||| || ||||| |||||||| || || || |||||| || | ||
Db        185 CCCATGGAGTGAATGCTCACGCACCTGCGGGGAGGGGCCTCCTACTCTCTGACGCCCTC 244

Qy        306 TTTGACTGGAAGGAATTGTGAAGGGCAGAACATTCGGTACAAGACATGCAGGCAATCATGA 365
                |||     | | ||  ||||||||      || || || ||||   ||||||||| |||     ||
Db        245 CCTGAGCAGCAAGAGCTGTGAAGGAACAAATATCCGATACAGAACATGCAGTAATGTGGA 304

Qy        366 CTGCCCTCCAGATGCAGAAGATTTCAGAGCCCAGCAGTGCTCAGCCTACAATGATGTCCA 425
              |||||| ||||| ||||   |||||| |||| ||||| ||||||||   | ||||||||| |
Db        305 CTGCCCACCAGAAGCAGGTGATTTCCGAGCTCAGCAATGCTCAGCTCATAATGATGTCAA 364

Qy        426 GTATCAGGGGCATTACTATGAATGGCTTCCACGATATAATGATCCTGCTGCCCCGTGTGC 485
              | | || || || |  |||||||||||||       | |||||| ||||      ||| ||| |
Db        365 GCACCATGGCCAGTTTTATGAATGGCTTCCTGTGTCTAATGACCCTGACAACCCATGTTC 424

Qy        486 ACTCAAGTGTCATGCACAAGGACAAAACTTGGTGGTGGAGCTGGCACCTAAGGTACTGGA 545
              ||||||||| || ||  |||||   || | |||| || || || |||||||||||   | ||
Db        425 ACTCAAGTGCCAAGCCAAAGGAACAACCCTGGTTGTTGAACTAGCACCTAAGGTCTTAGA 484

Qy        546 TGGAACTCGTTGCAACACGGACTCCTTGGACATGTGTATCAGTGGCATCTGTCAGGCAGT 605
              ||| || |||||| | || || || ||||| ||||| ||||||||   | || ||     ||
Db        485 TGGTACGCGTTGCTATACAGAATCTTTGGATATGTGCATCAGTGGTTTATGCCAAATTGT 544

Qy        606 GGGCTGCGATCGGCAACTGGGAAGCAATGCCAAGGA 641
               ||||||||||  || ||||||||||   | ||||||
Db        545 TGGCTGCGATCACCAGCTGGGAAGCACCGTCAAGGA 580


Search completed: March 22, 2003, 10:52:20
Job time : 4984 secs
```

```
                    GenCore version 5.1.4_p5_4578
             Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:          March 22, 2003, 06:25:33 ; Search time 125 Seconds
                                            (without alignments)
                                            12453.530 Million cell updates/sec

Title:           US-09-784-358-1
Perfect score:   5076
Sequence:        1 [illegible]........[illegible] 5076

Scoring table:   IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0

Searched:        441362 seqs, 1533380881 residues

Total number of hits satisfying chosen parameters:       882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Issued_Patents_NA:*
                 1:  /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
                 2:  /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
                 3:  /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
                 4:  /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
                 5:  /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
                 6:  /cgn2_6/ptodata/1/ina/backfiles1.seq:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                              SUMMARIES
                 %
Result           Query
  No.   Score    Match Length DB  ID                          Description
---------------------------------------------------------------------------

    1   733.2    14.4   1803  4   US-09-369-364A-20           Sequence 20, Appl
    2   109.8     2.2   2848  4   US-09-369-364A-4            Sequence 4, Appli
    3   101.4     2.0   3885  4   US-09-369-364A-16           Sequence 16, App.
    4    70.2     1.4   3675  4   US-09-930-872-3             Sequence 3, Appli
    5    70.2     1.4   4042  4   US-09-930-872-5             Sequence 5, Appl.
    6    67.4     1.3   3126  4   US-09-392-184-7             Sequence 7, Appli
    7    65.8     1.3   2114  4   US-09-130-491-7             Sequence 7, Appli
c   8      65     1.3   [illegible]  4   US-09-392-184-5      Sequence 5, Appli
    9    [illegible]  1.2   1642  4   US-09-369-364A-18       Sequence 18, Appl
   10    60.2     1.2   3218  4   US-09-369-364A-6            Sequence 6, Appli
   11    [illegible]  1.1   [illegible]  4   US-09-369-364A-8  Sequence 8, Appl
c  12      56     1.1   7218  4   US-08-232-463-14            Sequence 14, Appl
   13    54.6     1.1   2450  4   US-09-491-522-9             Sequence 9, Appli
   14    54.6     1.1   [illegible]  4   US-09-491-522-8      Sequence 8, Appli
   15    52.4     1.0   5804  4   US-09-369-364A-12           Sequence 12, Appl
   16    49.6     1.0   [illegible]  4   US-09-484-970B-58    Sequence 58, Appl
   17    49.6     1.0   [illegible]  4   [illegible]          Sequence 1, Appli
   18    49.6     1.0   4858  4   US-09-392-184-1             Sequence 1, Appli
   19      49     1.0   3002  4   US-09-369-364A-1            Sequence 1, Appli
   20      49     1.0   [illegible]  4   US-09-122-126B-1     Sequence 1, Appli
   21    47.6     0.9   2625  4   US-09-369-364A-14           Sequence 14, Appl
   22    44.4     0.9   [illegible]  4   US-09-369-364A-3     Sequence 3, Appli
   23    [illegible]  0.9   2130  4   US-09-491-522-10        Sequence 10, Appl
   24    44.4     0.9   3250  4   US-09-122-126B-14           Sequence 14, Appl
   25    43.8     0.9   2450  4   US-09-491-522-2             Sequence 2, Appli
   26    43.8     0.9   [illegible]  4   US-09-491-522-1      Sequence 1, Appli
c  27    41.6     0.8   2051  2   US-08-672-814D-1            Sequence 1, Appli
c  28    41.6     0.8   2051  4   US-09-333-696-1             Sequence 1, Appli
   29    41.2     0.8   2450  4   US-09-491-522-3             Sequence 3, Appli
c  30    36.6     0.7   1642  4   US-09-737-608B-24           Sequence 24, Appl
   31    36.2     0.7   1326  3   US-08-980-526-4             Sequence 4, Appli
   32    35.8     0.7   1470  4   US-09-099-041A-27           Sequence 27, Appl
   33    35.8     0.7   1470  4   US-09-245-281-27            Sequence 27, Appl
   34    35.8     0.7   1470  4   US-09-207-359B-27           Sequence 27, Appl
   35    35.8     0.7   [illegible]  [illegible]  [illegible]  Patent No. [illegible]
   36    35.8     0.7   2859  4   US-09-099-041A-9            Sequence 9, Appli
   37    35.8     0.7   2859  4   US-09-245-281-9             Sequence 9, Appli
   38    35.8     0.7   2859  4   US-09-207-359B-9            Sequence 9, Appli
   39    35.8     0.7   3080  4   US-09-099-041A-25           Sequence 25, Appl
   40    35.8     0.7   3080  4   US-09-245-281-25            Sequence 25, Appl
   41    35.8     0.7   3080  4   US-09-207-359B-25           Sequence 25, Appl
   42    35.8     0.7   3382  4   US-09-099-041A-7            Sequence 7, Appli
   43    35.8     0.7   3382  4   US-09-245-281-7             Sequence 7, Appli
   44    35.8     0.7   3382  4   US-09-207-359B-7            Sequence 7, Appli
   45    35.8     0.7   4302  4   US-09-245-281-38            Sequence 38, Appl

                              ALIGNMENTS

RESULT 1
US-09-369-364A-20
; Sequence 20, Application US/09369364A
; Patent No. 6391610
; GENERAL INFORMATION:
;  APPLICANT:  Apte, Suneel
;  APPLICANT:  Hurskainen, Tiina L.
;  APPLICANT:  Hirohata, Satoshi
;  TITLE OF INVENTION:  Nucleic Acids Encoding Zinc Metalloproteases
;  FILE REFERENCE:  [illegible]
;  CURRENT APPLICATION NUMBER:  US/09/369,364A
;  CURRENT FILING DATE:  1999-08-06
;  NUMBER OF SEQ ID NOS:  31
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 20
;   LENGTH: 1803
;   TYPE: DNA
;   ORGANISM: Homo sapiens ADAMTS R1
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (51)..(1634)
US-09-369-364A-20

  Query Match             14.4%;  Score 733.2;  DB 4;  Length 1803;
  Best Local Similarity   67.3%;  Pred. No. 2.2e-127;
  Matches 1035;  Conservative    0;  Mismatches  503;  Indels    0;  Gaps    0;

Qy        186 [illegible] 245
              [illegible]
Db        110 [illegible] 169

Qy        246 [illegible] 305
              [illegible]
Db        170 [illegible] 229

Qy        306 [illegible] 365
              [illegible]
Db        230 [illegible] 289

Qy        366 [illegible] 425
              [illegible]
Db        290 [illegible] 349

Qy        426 [illegible] 485
              [illegible]
Db        350 [illegible] 409

Qy        486 [illegible] 545
              [illegible]
Db        410 [illegible] 469
```

```
Qy        546 TGGAACTCGTTGCAACACGGACTCCTTGGACATGTGTATCAGTGGCATCTGTCAGGCAGT 605
              ||| || ||||||| | || || || ||||| ||||| ||||||||  | || ||    ||
Db        470 TGGTACGCGTTGCTATACAGAATCTTTGGATATGTGCATCAGTGGTTTATGCCAAATTGT 529

Qy        606 GGGCTGCGATGGGCAACTGGGAAGCAATGCCAAGGAGGACAACTGTGGAGTCTGTGCCGG 665
               |||||||||   || |||||||||||  | |||||| || |||||||| ||||   |||
Db        530 TGGCTGCGATCACCAGCTGGGAAGCACCGTCAAGGAAGATAACTGTGGGGTCTGCAACGG 589

Qy        666 CGATGGCTCCACCTGCAGGCTTGTACGGGGACAATCAAAGTCACACGTTTCTCCTGAAAA 725
               ||||| ||||||||| |||| || || || || ||  || || ||  | ||  |    ||
Db        590 AGATGGGTCCACCTGCCGGCTGGTCCGAGGGCAGTATAAATCCCAGCTCTCGGCAACCAA 649

Qy        726 AAGAGAAGAAAATGTAATTGCTGTTCCTTTGGCAAGTCGAAGTGTGAGAATTACAGTGAA 785
              |   || || | |||  ||||  |||| | | | |||| ||  | |  |  ||    | ||
Db        650 ATCGGATGATACTGTGGTTGCAATTCCCTATGGAAGTAGACATATTCGCCTTGTCTTAAA 709

Qy        786 ACGACCTGCCCACCTCTTTATTCAATCAAAAACACTTCAACGAAGCAAAGGAGAACACAG 845
              ||| ||||   ||| | | | | ||| | ||||| || || || |   ||||| ||| ||||
Db        710 AGGTCCTGATCACTTATATCTGGAAACCAAAACCCTCCAGGGACTAAAGGTGAAAACAG 769

Qy        846 CTTTAACAGCCCCGGGGTCTTTGTCGTAGAAAACACAACAGTGGAATTTCAGAGGGGCTC 905
                | | |  | | ||    ||  | || || |     |  ||||| || ||||      |
Db        770 TCTCAGCTCCACAGGAACTTTCCTTGTGGACAATTCTAGTGTGGACTTCCAGAAATTTCC 829

Qy        906 CGAGAGGCAAACTTTTAAGATTCCAGGACCTCTGATGGCTGATTTCATCTTCAAGACCAG 965
               || |   | |   | |  ||  | ||||| || |  || ||||||||  ||||||   |
Db        830 AGACAAAGAGATACTGAGAATGGCTGGACCACTCACAGCAGATTTCATTGTCAAGATTCG 889

Qy        966 GTACACTGCAGCCAAACACAGCCTGGTTCACTTCTTCTTTACCAGCCCATCAGTCATCA 1025
               || | |   ||   ||||    || |||||| |||| || || ||||||| || |
Db        890 TAACTCGGGCTCCGCTGACAGTACAGTCCAGTTCATCTTCTATCAACCCATCATCCACCG 949

Qy       1026 GTGGAGACAAACTGACTTCTTTCCCTGCACTGTGACGTGTGGAGGAGGTTATCAGCTCAA 1085
               |||||  | || || |||||||| ||| | |  |  |||||||||||||||||||| | 
Db        950 ATGGAGGGAGACGGATTTCTTTCCTTGCTCAGCAACCTGTGGAGGAGGTTATCAGCTGAC 1009

Qy       1086 TTCTGCTGAATGTGTGGATATCCGCTTGAAGAGGGTAGTTCCTGACCATTATTGTCACTA 1145
               || ||||| ||    ||| |       ||  | || ||| ||||||| || ||||||||
Db       1010 ATCGGCTGAGTGCTACGATCTGAGGAGCAACCGTGTGGTTGCTGACCAATACTGTCACTA 1069

Qy       1146 CTACCCTGAAAATGTAAAACCAAAACCAAAACTGAAGGAATGCAGCATGGATCCCTGCCC 1205
               ||||| || ||    |||||| ||||| || ||  |||| ||||   ||||||| || ||
Db       1070 TTACCCAGAGAACATCAAACCCAAACCCAAGCTTCAGGAGTGCAACTTGGATCCTTGTCC 1129

Qy       1206 ATCAAGTGATGGATTTAAAGAGATAATGCCCTATGACCACTTCCAACCTCTTCCTCGCTG 1265
              |   |||||  ||||  ||  |||| ||||| ||||||| || ||| || |||||||| ||
Db       1130 AGCCAGTGACGGATACAAGCAGATCATGCCTTATGACCTCTACCATCCCCTTCCTCGGTG 1189

Qy       1266 GGAACATAATCCTTGGACTGCATGTTCCGTGTCCTGTGGAGGAGGGATTCAGAGACGGAG 1325
              |||    |  || ||||| || || |||   || ||||| || || || ||||| |||  
Db       1190 GGAGGCCACCCCATGGACCGCGTGCTCCTCCTCGTGTGGGGGGGGCATCCAGAGCCGGGC 1249

Qy       1326 CTTTGTGTGTGTAGAGGAATCCATGCATGGAGAGATATTGCAGGTGGAAGAATGGAAGTG 1385
                ||    |||| |||||   ||| || ||     |       ||||||||| ||||| ||
Db       1250 AGTTTCCTGTGTGGAGGAGGACATCCAGGGGCATGTCACTTCAGTGGAAGAGTGGAAATG 1309

Qy       1386 CATGTACGCACCCAAACCCAAGGTTATGCAAACTTGTAATCTGTTTGATTGCCCCAAGTG 1445
              ||||||| | ||| ||        |     | |     | |  ||||| ||||| || ||
Db       1310 CATGTACACCCCCTAAGATGCCCATCGCGCAGCCCTGCAACATTTTGACTGCCCTAAATG 1369

Qy       1446 GATTGCCATGGAGTGGTCTCAGTGCACAGTGACTTGTGGCCGAGGGTTACGGTACCGGGT 1505
              | | ||   ||||||||||| ||||||||||||| |||||||  ||  |  | ||||| ||
Db       1370 GCTGGCACAGGAGTGGTCTCCGTGCACAGTGACGTGTGGCCAGGGCCTCAGATACCGTGT 1429

Qy       1506 TGTTCTGTGTATTAACCACCGCGGAGAGCATGTTGGGGGCTGCAATCCACAACTGAAGTT 1565
               || || || ||  |||| || |||  |||    || ||||| |  ||| ||   |||  
Db       1430 GGTCCTCTGCATCGACCATCGAGGAATGCACACAGGAGGCTGTAGCCCAAAAACAAAGCC 1489

Qy       1566 ACACATCAAAGAAGAATGTGTCATTCCCATCCCGTCTTATAAACCAAAAGAAAAAAGTCC 1625
               ||||| |||||| |||||  || | |||| ||  | ||||||||| ||||| |||  |||
Db       1490 CCACATAAAAGAGGAATGCATCGTACCCACTCCCTCCTATAAACCCAAAGAGAAACTTCC 1549
```

```
Qy       1626 AGTGGAAGCAAAATTGCCTTGGCTGAAACAAGCACAAGAACTAGAAGAGACCAGAATAGC 1685
              ||| || || || ||||| ||| | |||||||| ||||| ||||||||          | 
Db       1550 AGTCGAGGCCAAGTTGCCATGGTTCAAACAAGCTCAAGAGCTAGAAGAAGGAGCTGCTGT 1609

Qy       1686 AACAGAAGAACCAACGTTCATTCCAGAACCCTGGTCAG 1723
                |||| || ||  |||   ||  | || |   | | |
Db       1610 GTCAGAGGAGCCCTCGTAAGTTGTAAAAGCACAGACTG 1647


RESULT 2
US-09-369-364A-4
; Sequence 4, Application US/09369364A
; Patent No. 6391610
; GENERAL INFORMATION:
;  APPLICANT:  Apte, Suneel
;  APPLICANT:  Hurskainen, Tiina L.
;  APPLICANT:  Hirohata, Satoshi
;  TITLE OF INVENTION:  Nucleic Acids Encoding Zinc Metalloproteases
;  FILE REFERENCE:  26473/1007/10-30-00
;  CURRENT APPLICATION NUMBER: US/09/369,364A
;  CURRENT FILING DATE:  1999-08-06
;  NUMBER OF SEQ ID NOS:  31
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 4
;   LENGTH: 2848
;   TYPE: DNA
;   ORGANISM: Homo sapiens ADAMTS-6
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (22)..(2601)
;   NAME/KEY: misc_feature
;   LOCATION: (1369)
;   OTHER INFORMATION: n = C
;   NAME/KEY: misc_feature
;   LOCATION: (1620)
;   OTHER INFORMATION: n=C
US-09-369-364A-4

  Query Match             2.2%;  Score 109.8;  DB 4;  Length 2848;
  Best Local Similarity   54.0%;  Pred. No. 1.6e-24;
  Matches  259;  Conservative    0;  Mismatches  203;  Indels   18;  Gaps    1;

Qy        219 CAAAGATGGCAACTGGGATGCTTGGGGCGACTGGAGTGACTGCTCCCGGACCTGTGGGGG 278
              || ||||||   ||||| | | |||         ||| | || |||   | ||||||| |||||
Db       1545 CATAGATGGGGGCTGGGTCCCTGGTCACTATGGGAGACTCCAGCAGGACCTGCGGGGG 1604

Qy        279 AGGAGCATCATATTCTCTGCGGAGATGTTTGACT------------------GGAAGGAA 320
              ||| |  || |  || ||  |     ||   | |                   | ||   |
Db       1605 AGGCGTNTCCTCATCCCTAAGACACTGTGACAGTCCAGCACCTTCGGAGGTGGAAAAATA 1664

Qy        321 TTGTGAAGGGCAGAACATTCGGTACAAGACATGCAGCAATCATGACTGCCCTCCAGATGC 380
              |||      ||| | |  |  ||||      | || | ||    ||     ||||||    | | |
Db       1665 TTGCCTTGGGAAAGGAAACGGTATCGCTCCTGTAACACAGATCCATGCCCTTTGGGTTC 1724

Qy        381 AGAAGATTTCAGAGCCCAGCAGTGCTCAGCCTACAATGATGTCCAGTATCAGGGGCATTA 440
                 ||||||  |||   | |||||  ||| ||   |  ||   |  |  |  ||    ||
Db       1725 CCGAGATTTTCGAGAGAAACAGTGTGCAGACTTTGACAATATGCCTTTCCGAGGAAAGTA 1784

Qy        441 CTATGAATGGCTTCCACGATATAATGATCCTGCTGCCCCGTGTGCACTCAAGTGTCATGC 500
               ||| |  ||   ||      |     |   |     || |||||| | || ||    ||
Db       1785 TTATAACTGGAAACCCTATACTGGAGGTGGGGTAAAACCTTGTGCATTAAACTGCTTGGC 1844

Qy        501 ACAAGGACAAAACTTGGTGGTGGAGCTGGCACCTAAGGTACTGGATGGAACTCGTTGCAA 560
                ||||    || ||       ||    || |||  |||    ||||| || |  |||||
Db       1845 TGAAGGTTATAATTTCTACACTGAACGTGCTCCTGCGGTGATCGATGGGACCCAGTGCAA 1904

Qy        561 CACGGACTCCTTGGACATGTGTATCAGTGGCATCTGTCAGGCAGTGGGCTGCGATCGGCA 620
               |||| ||   |||| || || |||| |||      ||   ||     || ||||| |||
Db       1905 TGCCGATTCACTGGATATCTGCATCAATGGAGAATGCAAGCACGTAGGCTGTGATAATAT 1964

Qy        621 ACTGGGAAGCAATGCCAAGGAGGACAACTGTGGAGTCTGTGCCGGCGATGGCTCCACCTG 680
               |||||     |||| | ||| || |   ||| ||||||||| || |   ||    ||| ||
Db       1965 TTTGGGATCTGATGCTAGGGAAGATAGATGTCCGAGTCTGTGGAGGGGGCGGAAGCACATG 2024
```

```
RESULT 3
US-09-369-364A-16
; Sequence 16, Application US/09369364A
; Patent No. 6391610
; GENERAL INFORMATION:
;  APPLICANT:  Apte, Suneel
;  APPLICANT:  Hurskainen, Tiina L.
;  APPLICANT:  Hirohata, Satoshi
;  TITLE OF INVENTION:  Nucleic Acids Encoding Zinc Metalloproteases
;  FILE REFERENCE:  26473/4007/10-30-00
;  CURRENT APPLICATION NUMBER: US/09/369,364A
;  CURRENT FILING DATE:  1999-08-06
;  NUMBER OF SEQ ID NOS:  31
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 16
;   LENGTH: 3885
;   TYPE: DNA
;   ORGANISM: Homo sapiens ADAMTS-16
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(3246)
;   NAME/KEY: misc_feature
;   LOCATION: (3877)
;   OTHER INFORMATION: n=A
;   NAME/KEY: misc_feature
;   LOCATION: (3882)
;   OTHER INFORMATION: n=G
US-09-369-364A-16

  Query Match             2.0%;  Score 101.4;  DB 4;  Length 3885;
  Best Local Similarity   53.0%;  Pred. No. 1.1e-21;
  Matches  253;  Conservative    0;  Mismatches  206;  Indels   18;  Gaps    1;

Qy        223 GATGGCAACTGGGATGCTTGGGCCGACTGGAGTGACTGCTCCCGGACCTGTGGGGGAGGA 282
Db       1570 GACGGAGCCTGGGGGCCGTGGACTCCATGGGGCGACTGCAGCCGGACCTGTGGCGGCGGC 1629

Qy        283 GCATCATATTC                       CTGTGGAGATGTTTGACTGGAAGGAATTGT 324
Db       1630 [illegible] 1689

Qy        325 GAAGGGCAGAACATT[illegible]AGATGCAGAA 384
Db       1690 [illegible]AAC[illegible]ATGACTG[illegible]AG 1749

Qy        385 GATTTCAGAGCCCAGCAGTGCTCAGCCTACAATGATGTCCAGTATCAGGGGCATTACTAT 444
Db       1750 GACTTCAGAGAAGTGCAGTGTGCTGAATTTGA[illegible]GAAATTCTAC 1809

Qy        445 GAATGGCTTCCACGATATAATGATCCTGCTGCCCCGTGTGCACTTAAGTGTCATGCACAA 504
Db       1810 AAGTGGAAAA[illegible] 1869

Qy        505 GGACAAAACTTGGTTGTGGAGCTGGCACCTAAGGTACTGGATGGAACTCGTTGCAATACA 564
Db       1870 [illegible] 1929

Qy        565 GACTCCTTGGACATG[illegible]AACTGT 624
Db       1930 GACAGGTGGACATT[illegible]GACCGAGTCCTG 1989

Qy        625 GGAAGCAATGCCAAGGAGGACAACTGTGGAGTCTGTGCCGGCGATGGCTCCACCTGC 681
Db       1990 [illegible]AAGTGCCGAGTGTGTGGCGGTGACGGCAGTGCCTGC 2046


RESULT 4
US-09-930-872-3
; Sequence 3, Application US/09930872
; Patent No. 6448388
; GENERAL INFORMATION:
;  APPLICANT: Friddle, Carl Johan
;  APPLICANT:  Hilbun, Erin
;  TITLE OF INVENTION: No. 6448388el Human Proteases and Polynucleotides Encoding Th
;  FILE REFERENCE: LEX-0219-USA
;  CURRENT APPLICATION NUMBER: US/09/930,872
;  CURRENT FILING DATE:  2001-08-14
;  PRIOR APPLICATION NUMBER: US 60/225,852
;  PRIOR FILING DATE: 2000-08-16
;  NUMBER OF SEQ ID NOS: 5
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 3675
;   TYPE: DNA
;   ORGANISM: homo sapiens
US-09-930-872-3

  Query Match             1.4%;  Score 70.2;  DB 4;  Length 3675;
  Best Local Similarity   51.8%;  Pred. No. 1.4e-11;
  Matches  184;  Conservative    0;  Mismatches  168;  Indels    3;  Gaps    1;

Qy        346 AAGACATGCAGCAATCATGA[illegible]AGACAGCAGAAGATTTCAGAG[illegible] 405
Db       1897 AAGCTCTGCAACAGTCAGAAATG[illegible]GACATTGTTGACTT[illegible]CAGTGT 1956

Qy        406 TCAGCCTACAATGATG[illegible]AATACAGGGCATTACTATGAA[illegible]CCACTATATAA 464
Db       1957 [illegible]CAGACAA[illegible]AGACGATTCAGA[illegible]TACTAAT[illegible]AATTACACTCAA 2016

Qy        465 [illegible]AAGTGTCATGCACAAGGACAAAACTTGGTTGTGG 522
Db       2017 GTAGAAGATCAGGACTTA[illegible]AAACTCTAC[illegible]AAGGATTTGAT[illegible]TTCTTT 2076

Qy        523 GAGCTGGCACCTAAGGTACTGGATGGAACTCGTTGCAATACAGGACTCCTTGGACATGTGT 582
Db       2077 TCTTTGTCAAATAAAGTCAAAGATGGGACTCCATGCTGGAGGATAGCCGTAATGTTTGT 2136

Qy        583 ATCAGTGGCATCTGTCAGGCAGTGGGCTGCGATCGGCAACTGGGAAGCAATGCCAAGGAG 642
Db       2137 ATAGATGGGATATGTGAGAGAGTTGGATCTGACAATGTCCTTGGATCTGATGCTGTTGAA 2196

Qy        643 GACAACTGTGGAGTCTGTGCCGGCGATGGCTCCACCTGCAGGCTTGTACGGGGAC 697
Db       2197 GACGT[illegible]GTAA[illegible]AATAA[illegible]ATT[illegible] 2251


RESULT 5
US-09-930-872-5
; Sequence 5, Application US/09930872
; Patent No. 6448388
; GENERAL INFORMATION:
;  APPLICANT: Friddle, Carl Johan
;  APPLICANT:  Hilbun, Erin
;  TITLE OF INVENTION: No. 6448388el Human Proteases and Polynucleotides Encoding the
;  FILE REFERENCE: LEX-0219-USA
;  CURRENT APPLICATION NUMBER: US/09/930,872
;  CURRENT FILING DATE:  2001-08-14
;  PRIOR APPLICATION NUMBER: US 60/225,852
;  PRIOR FILING DATE: 2000-08-16
;  NUMBER OF SEQ ID NOS: 5
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
;   LENGTH: 4042
;   TYPE: DNA
;   ORGANISM: homo sapiens
US-09-930-872-5

  Query Match             1.4%;  Score 70.2;  DB 4;  Length 4042;
  Best Local Similarity   51.8%;  Pred. No. 1.6e-11;
  Matches  184;  Conservative    0;  Mismatches  169;  Indels    3;  Gaps    1;

Qy        346 AAGACATGCAGCAATCATGA[illegible]CAGATGCAGAAGATTTCAGAGCCCAGCAGTGC 405
Db       1995 AAGCTCTGCAACAGTCAGAAATG[illegible]GGGACAGTGTTGACTTCGGTGCTGCTCAGTGT 2054
```

```
Qy       406 TCAGCCTACAATGATGTCCAGTATCAGGGGCATTACTATGAATGGCTTCCACGATATAA- 464
                ||    |||           |    ||||  ||||  | |||  ||        ||
Db      2055 GCCGAGCACAACAGCAGACGATTCAGAGGGCGGCACTACAAGTGGAAGCCTTACACTCAA 2114

Qy       465 --TGATCCTGCTGCCCCGTGTGCACTCAAGTGTCATGCACAAGGACAAAACTTGGTGGTG 522
                  |   |   | |   ||   ||:| | |||   ||| :||||    | ||  |  |
Db      2115 GTAGAAGATCAGGACTTATGCAAACTCTACTGTATCGCAGAAGGATTTGATTTCTTCTTT 2174

Qy       523 GAGCTGGCACCTAAGGTACTGGATGGAACTCGTTGCAACACGGACTCCTTGGACATGTGT 582
                 || ||  ||| ||     ||||| ||||  |||     |||   |    |  | |||
Db      2175 TCTTTGTCAAATAAAGTCAAAGATGGGACTCCATGCTCGGAGGATAGCCGTAATGTTTGT 2234

Qy       583 ATCAGTGGCATCTGTCAGGCAGTGGGCTGCCGATCGGCAACTGGGAAGCAATGCCAAGGAG 642
             ||    ||| || ||| ||   ||| || || ||          || |||    ||||    ||
Db      2235 ATACATGCGATATCTGACACACTTGGATGTGACAATCTCCTTGCATCTCATGCTCTTGAA 2294

Qy       643 GACAACTGTGGAGTCTGTGCCGGCCGATGGCTCCACCTGCAGCCTTGTACCGGGAC 697
             |||  ||||||  ||  ||   ||   ||  |||  ||||||| | ||     |||| |
Db      2295 GACGTCTGTGGGGTGTGTAACGGGAATAACTCAGCCTGCACGATTCACAGGGGTC 2349


RESULT 6
US-09-392-184-7/c
; Sequence 7, Application US/09392184
; Patent No. 6395889
; GENERAL INFORMATION:
;  APPLICANT:  Robison, Keith E.
;  TITLE OF INVENTION:  NUCLEIC ACID MOLECULES ENCODING HUMAN
;  TITLE OF INVENTION:  PROTEASE HOMOLOGS
;  FILE REFERENCE:  5800-55
;  CURRENT APPLICATION NUMBER: US/09/392,184
;  CURRENT FILING DATE:  1999-09-09
;  NUMBER OF SEQ ID NOS:  33
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 7
;    LENGTH: 3126
;    TYPE: DNA
;    ORGANISM: Homo sapiens
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (1)...(3126)
;    OTHER INFORMATION: reprolysin (ADAM family of metalloprotease)
US-09-392-184-7

  Query Match              1.3%;  Score 67.4;  DB 4;  Length 3126;
  Best Local Similarity   51.3%;  Pred. No. 1e-10;
  Matches  257;  Conservative    0;  Mismatches  211;  Indels   33;  Gaps    3;

Qy       222 AGATGGCAACTGGGATGCTTGGGGCGACTGGAGTGACTGCTCCGGGACCTGTGGGGGAGG 281
             ||||||   |||||    | |||||   |||| | || || || ||||||||||| |||||
Db      1750 AGATGGAGGCTGGGCACCGTGGGGACCCTGGGGAGAATGTTCTCGGACCTGTGGAGGAGG 1691

Qy       282 AGCATCATATTCTCTGCGGAGATGTTTGACT                    GGAAGGAATTG 323
             || |   | |||   ||    ||   |                      |||||  | ||
Db      1690 AGTACAGTTTTCACACCGTGAGTGCAAGGACCCCGAGCCTCAGAATGGAGGAAGATACTG 1631

Qy       324 TGAAGGGCAGAACATCCCCTACAACACATCCAGCAATCATCACTGCCCTCCAGATGCAGA 383
                 || | ||         |||| || |||||  ||    | ||  |||||| || || |    |
Db      1630 CCTGGGTCGGAGAGCCAAGTACCAGTCATGCCACACGGAGGAATGCCCCCCTGACGGGAA 1571

Qy       384 AGATTTCAGAGCCCAGCAGTGCTCA---------GCCTACAATGATGTCCAGTATCAGGG 434
             |   ||||| |  ||||||||                ||||||||| |     |     | ||
Db      1570 AAGCTTCAGGGAGCAGCAGTGTGAGAAGTATAATGCCTACAATTACACTGACATGGACGG 1511

Qy       435 GCATTACTATGAATGGCTTCCACGATATAATGATCCTGCTGCCCCGTGTGCACTCAAG   492
             | ||  |    | ||| | ||    |||  ||       |  ||||  |          ||||
Db      1510 GAATCTCCTGCAGTGGGTCCCCAAGTATGCTGGGGTGTCCCCCCGGGACCGCTGCAAGTT 1451

Qy       493 ----TGTCATGCACAAGGACAAAACTTGGTGGTGGAGCTGGCACCTAAGGTACTGGATGG 548
                 || |  || |  ||      | |  | |     | | | |    | |||||  | |||||
Db      1450 GTTCTGCCGAGGCCGGGGAGGACCGAGTTCAAAGTGTTCGAGCCCAAGCTGATTGATGC 1391
```

```
Qy       549 AACTCGTTGCAACACGGACTCCTTGGACATGTGTATCAGTGGCATCTGTCAGGCAGTGGG 608
               || |  ||      | ||  |  || ||| ||| || |||||    |||        |  ||
Db      1390 CACCCTGTGTGGGCCAGAAACACTGGCCATCTGTGTCGTGGCCAGTGTGTCAAGGCCGG 1331

Qy       609 CTGCCGATCGGCAACTGGGAAGCAATGCCAAGGAGGACAACTGTGGAGTCTGTGCCGGCGA 668
             ||| || |      |||        |     |||  |||||| || || || |||| ||| |
Db      1330 CTGTGACCATGTGGTGGACTCGCCTCGGAAGCTGGACAAATGCGGGGTGTGTGGGGGCAA 1271

Qy       669 TGGCTCCACCTGCAGGCTTGT 689
               |||  | ||||||||    ||
Db      1270 AGGCAACTCCTGCAGGAAGGT 1250


RESULT 7
US-09-130-491-7
; Sequence 7, Application US/09130491
; Patent No. 6416974
; GENERAL INFORMATION:
;  APPLICANT: Holtzman, Douglas A.
;  APPLICANT: Goodearl, Andrew D.J.
;  TITLE OF INVENTION: TANGO 71, TANGO 73, TANGO 74, TANGO 76, AND TANGO 83
;  FILE REFERENCE: 09404/041001
;  CURRENT APPLICATION NUMBER: US/09/130,491
;  CURRENT FILING DATE: 1998-08-07
;  EARLIER APPLICATION NUMBER: US 60/058,108
;  EARLIER FILING DATE: 1997-09-05
;  EARLIER APPLICATION NUMBER: US 60/054,961
;  EARLIER FILING DATE: 1997-08-06
;  NUMBER OF SEQ ID NOS: 16
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 7
;    LENGTH: 2114
;    TYPE: DNA
;    ORGANISM: Rattus rattus
;    FEATURE:
;    NAME/KEY: CDS
;    LOCATION: (3)...(1445)
US-09-130-491-7

  Query Match              1.3%;  Score 65.8;  DB 4;  Length 2114;
  Best Local Similarity   51.1%;  Pred. No. 2.6e-10;
  Matches  256;  Conservative    0;  Mismatches  212;  Indels   33;  Gaps    3;

Qy       222 AGATGGCAACTGGGATGCTTGGGGCGACTGGAGTGACTGCTCCGGGACCTGTGGGGGAGG 281
             ||||||  |||||| | | |||||   |||| |  | ||  || || |||||||| |||||
Db       353 AGATGGAGACTGGGGTCCCTGGGGACCCTGGGGACAATGTTCTCCCACCTGTGGTGGAGG 412

Qy       282 AGCATCATATTCTCTGCGGAGATGTTTGACT------ -----------GGAAGGAATTG 323
                |   | |||    ||     |||   | |                    |||||   |||
Db       413 GATACACTTTTCGAACCGTCAGTGTGATAATCCAGCACCTCAGAATGGAGGAAGATTTTG 472

Qy       324 TGAAGGGCAGAACATCCCCTACAACACATCCAGCAATCATCACTGCCCTCCAGATGCAGA 383
                 ||  |||    |    |||| |   | ||||   |     | ||  ||  || ||| ||| | |
Db       473 GCTGGGAGAGAGAGTCAAGTACCAATCTTGCAAGACAGAGGAATGTCCACCAAATGGAAA 532

Qy       384 AGATTTCAGAGCCCAGCAGTGCTCA---------GCCTACAATGATGTCCAGTATCAGGG 434
             |     ||||| |   ||||||||     |              |||||||    |     |     | ||
Db       533 AACTTCTACCTAGCAGCAGCACTTCAAAAACATAATCTTCACAACCACAGGGAGCTGGATGG 592

Qy       435 GCATTACTATGAATGGCTTCCACGATATAATGAT------CCTGCTGCCCCGTGTGCACT 488
             | ||| |   | | ||| | ||     |||     |        ||     | ||  ||     |||
Db       593 GAATTTCCTTCAGTGGGTCCCCAAATACTCAGGAGTGTCCCCCCGAGACCGATGCAAACT 652

Qy       489 CAAGTGTCATGCACAAGGACAAAACTTGGTGGTGGAGCTGGCACCTAAGGTACTGGATGG 548
                 ||     || |  ||      |     | |     | | | | | | |||||||   |  |||||
Db       653 GTTTGCAGAGGCCGTGGGAGGAGTGAGTTCAAAGTGTTTGAAACTAAGGTGATCGATGG 712

Qy       549 AACTCGTTGCAACACGGACTCCTTGGACATGTGTATCAGTGGCATCTGTCAGGCAGTGGG 608
              ||||  |||     ||||  |   ||| ||| ||| |  | ||     ||        |  ||
Db       713 CACTCTGTGCGGACCGGATACTCTGGCCATCTGTGTCGTGGGACAGTGCGTTAAGGCTGG 772

Qy       609 CTGCCGATCGGCAACTGGGAAGCAATGCCAAGGAGGACAACTGTGGAGTCTGTGCCGGCGA 668
             ||| || |       ||          |     |||  |||||| ||||| || |||| ||| |
Db        [illegible]

Qy        [illegible] 689
              [illegible]
Db        [illegible] 853


RESULT 8
US-09-392-184-5/c
; Sequence 5, Application US/09392184
; Patent No. [illegible]
; GENERAL INFORMATION:
;  APPLICANT:  Robison, Keith E.
;  TITLE OF INVENTION:  NUCLEIC ACID MOLECULES ENCODING HUMAN
;  TITLE OF INVENTION:  PROTEASE HOMOLOGS
;  FILE REFERENCE:  5800-55
;  CURRENT APPLICATION NUMBER: US/09/392,184
;  CURRENT FILING DATE:  1999-09-09
;  NUMBER OF SEQ ID NOS:  33
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 5
;   LENGTH: 5357
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(5357)
;   OTHER INFORMATION: reprolysin (ADAM family of metalloprotease)
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(5357)
;   OTHER INFORMATION: n = A,T,C or G
US-09-392-184-5

  Query Match             1.3%;  Score 65;  DB 4;  Length 5357;
  Best Local Similarity   52.4%;  Pred. No. [illegible];
  Matches  196;  Conservative    0;  Mismatches  170;  Indels    9;  Gaps    2;

Qy        313 [illegible] 372
              [illegible]
Db       3508 [illegible] 3567

Qy        373 [illegible] 432
              [illegible]
Db       3568 [illegible]

Qy        433 [illegible]
              [illegible]
Db        [illegible]

Qy        [illegible]
              [illegible]
Db        [illegible]

Qy        553 [illegible]
              [illegible]
Db        [illegible]

Qy        [illegible] 666
              [illegible]
Db        [illegible]

Qy        667 [illegible]
              [illegible]
Db       3241 [illegible]


RESULT 9
US-09-369-364A-18
; Sequence 18, Application US/09369364A
; Patent No. 6391610
; GENERAL INFORMATION:
;  APPLICANT:  Apte, Suneel
;  APPLICANT:  Hurskainen, Tiina L.
;  APPLICANT:  Hirohata, Satoshi
;  TITLE OF INVENTION:  Nucleic Acids Encoding Zinc Metalloproteases
;  FILE REFERENCE:  26473/4007/[illegible]
;  CURRENT APPLICATION NUMBER: US/09/369,364A
;  CURRENT FILING DATE:  1999-08-06
;  NUMBER OF SEQ ID NOS:  41
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 18
;   LENGTH: 1642
;   TYPE: DNA
;   ORGANISM: Mus musculus ADAMTS 10
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (2)..(1351)
;   NAME/KEY: misc_feature
;   LOCATION: (354)
;   OTHER INFORMATION: n=C
;   NAME/KEY: misc_feature
;   LOCATION: (338)
;   OTHER INFORMATION: n=C
;   NAME/KEY: misc_feature
;   LOCATION: (1511)
;   OTHER INFORMATION: n=A, T, G, C, or T
;   NAME/KEY: misc_feature
;   LOCATION: (1553)
;   OTHER INFORMATION: n=A, T, G, C, or T
;   NAME/KEY: misc_feature
;   LOCATION: (1584)
;   OTHER INFORMATION: n=A, T, G, C, or T
;   NAME/KEY: misc_feature
;   LOCATION: (1616)
;   OTHER INFORMATION: n=A, T, G, C, or T
;   NAME/KEY: misc_feature
;   LOCATION: (1642)
;   OTHER INFORMATION: n=A, T, G, C, or T
US-09-369-364A-18

  Query Match             1.2%;  Score 62.4;  DB 4;  Length 1642;
  Best Local Similarity   63.2%;  Pred. No. 2.7e-09;
  Matches   96;  Conservative    0;  Mismatches   56;  Indels    0;  Gaps    0;

Qy        529 [illegible] 588
              [illegible]
Db          2 [illegible] 61

Qy        589 [illegible] 648
              [illegible]
Db         62 [illegible] 121

Qy        649 [illegible] 680
              [illegible]
Db        122 [illegible] 153


RESULT 10
US-09-369-364A-6
; Sequence 6, Application US/09369364A
; Patent No. 6391610
; GENERAL INFORMATION:
;  APPLICANT:  Apte, Suneel
;  APPLICANT:  Hurskainen, Tiina L.
;  APPLICANT:  Hirohata, Satoshi
;  TITLE OF INVENTION:  Nucleic Acids Encoding Zinc Metalloproteases
;  FILE REFERENCE:  26473/4007/[illegible]
;  CURRENT APPLICATION NUMBER: US/09/369,364A
;  CURRENT FILING DATE:  1999-08-06
;  NUMBER OF SEQ ID NOS:  41
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 6
;   LENGTH: 3218
;   TYPE: DNA
;   ORGANISM: Homo sapiens ADAMTS-7
;   FEATURE:
```

```
;    NAME/KEY: CDS
;    LOCATION: (13)..(3003)
US-09-369-364A-6

  Query Match             1.2%;  Score 60.2;  DB 1;  Length 3218;
  Best Local Similarity   51.5%;  Pred. No. 2.3e-08;
  Matches  193;  Conservative    0;  Mismatches  173;  Indels    9;  Gaps    2;

Qy        313 GGAACCAATTCTCAAGGCCAGAACATTCGCTACAACACATGCAGCAATCATGACTGCCCT 372
              || ||   | ||||  ||   ||   ||   || | |        |||| |    || | |||||||
Db       1732 GGCAGATACTGTGTGGGTGAGCGCAAGCGCTTCCGCCTCTGCAACCTGCAGGCCTGCCCT 1791

Qy        373 CCAGATGCAGAAGATTTCAGAGCCCAGCAGTGCTCAGCCTACAATGATGTCCAGTATCAG 432
               | |              ||| |    |    ||||||     ||    | | | | |  ||   ||
Db       1792 GCTGGCCGCCCCTCCTTCCGCCACGTCCAGTGCAGCCACTTTGACGCTATGCTCTACAAG 1851

Qy        433 GGGCATTACTATGAATGGCTTCCACGATATAATCATCCTGCTGCCCCGTGTCACTCAAC 492
              || ||      |    |||| | .|          |||||      |      ||| || |   ||   |
Db       1852 GGCCAGCTGCACACATGGGTGCCCGTGGTCAATGA---CGTGAACCCCTGCGAGCTGCAC 1908

Qy        493 TGTCATGCACAAGGACAAAACTTGGTGGTGGAGCTGGCACCTAAGGTACTCCATCCAACT 552
              || |   |          |  ||||  |     | |||||               | ||||| ||
Db       1909 TGCCGGCCCGGCGAATGAGTACTTTGCCAAGAAGCTGCGGGACGCCTGTGTCGATGGCACC 1968

Qy        553 CGTTGCAACACG------GACTCCTTGGACATGTGTATCAGTGGCATCTGTCAGGCAGTG 606
               |   ||| ||   |        | |    |   |||| |  || ||||   ||||||||| ||      |||
Db       1969 CCCTGCTACCAGGTTCGAGCCAGCCGGGACCTCTGCATCAACCGCATCTGTAAGAACGTG 2028

Qy        607 GGCTGCGATCGGCAACTGGGAAGCAATGCCAAGGAGGACAACTGTGGAGTCTGTGCCGGC 666
              ||||| ||        |  |  |      |   ||| | |||||||  |||||| || ||    ||||
Db       2029 GGCTGTGACTTCGAGATTGACTCCGGTGCTATGGAGGACGGCTGTGGTGTGTGCCACGGC 2088

Qy        667 GATGGCTCCACCTGC 681
                | |||||||||||
Db       2089 AACGGCTCCACCTGC 2103


RESULT 11
US-09-369-364A-8
; Sequence 8, Application US/09369364A
; Patent No. 6391610
; GENERAL INFORMATION:
;  APPLICANT:  Apte, Suneel
;  APPLICANT:  Hurskainen, Tiina L.
;  APPLICANT:  Hirohata, Satoshi
;  TITLE OF INVENTION:  Nucleic Acids Encoding Zinc Metalloproteases
;  FILE REFERENCE:  26473/4007/10-30-00
;  CURRENT APPLICATION NUMBER: US/09/369,364A
;  CURRENT FILING DATE:  1999-08-06
;  NUMBER OF SEQ ID NOS:  31
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 8
;   LENGTH: 3638
;   TYPE: DNA
;   ORGANISM: Mus musculus ADAMTS 8
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (278)..(2992)
;   NAME/KEY: misc_feature
;   LOCATION: (3636)
;   OTHER INFORMATION: n    T
US-09-369-364A-8

  Query Match             1.1%;  Score 57.6;  DB 4;  Length 3638;
  Best Local Similarity   50.2%;  Pred. No. 1.8e-07;
  Matches  249;  Conservative    0;  Mismatches  214;  Indels   33;  Gaps    4;

Qy        222 AGATGGCAACTGGGATGCTTGGGGCGACTGGAGTGACTGCTCCCGGACCTGTGGGGGAGG 281
              ||||||  |||||| | | |||  |    |||| |   | || || || |||||||| |||||
Db       1900 AGATGGAGACTGGGGTCCCTGGAGACCCTGGGGACAATGTTCTCGCACCTGTGGTGGAGG 1959

Qy        282 AGCATCATATTCTCTGGGGAGATGTTTGACT                ----GGAAGGAATTG 323
                  | ||  |.     ||   ||||   | |                        |||||   |||
Db       1960 CATACAATTCCCAACCCTGAATGTGATAATCCAATGCCTCAGAATCGAGGAACATTTTG 2019

Qy        324 TGAAGGGCAGAACATTCGGTACAAGACATGCAGCAATCATGACTGCCCTCCAGATGCAGA 383
                   ||   | |    |    |||| |   |||||| ||     | ||  || || ||| | | | |
Db       2020 CCTGGGTGAAAGACTCAACTACCAATCATGAAACAGAGCAATGTTCATTAAAGGAAA 2079

Qy        384 AGATTTCAGAGCCCAGCAGTGCTCA---------GCCTACAATGATGTCCAGTATCAGGG 434
               |    ||| |  |   ||||||||                 ||||||||    |        |      | ||
Db       2080 AAGCTTCCGCCAGCACCAGTGTCAGAAATATAATGCCTACAACCACACTGACCTGGATGG 2139

Qy        435 GCATTACTATGAATCGCTTCCACCATATAATC------ATCCTGCTGCCCCGTGTGCACT 488
               | ||| |      | |||| | ||      |||     |             ||      | ||   ||        ||
Db       2140 GAATTTCCTGCAGTCGGTCCCCAAGTATTCAGGAGTGTCCCCCGAGACCGATGCAAGCT 2199

Qy        489 CAAGTGTCATGCACAAGGACAAAACTTGGTGGTGGAGCTGGCACCTAAGGTACTGGATGG 548
                   ||      || |  ||      |      | |       | | | | | | |||||||  | |||||
Db       2200 GTTTTGCAGAGCCCTTGGGAGGAGTGAGTTCAAAGTGTTTGAAGCTAAGGTGATCCATCC 2259

Qy        549 AACTCGTTGCAACACGGACTCCTTGGACATGTGTATCAGTGGCATCTGTCAGGCAGTGGG 608
                 ||||  ||       ||||  |  ||   ||| ||   || | | ||      |||        |  ||
Db       2260 CACTCTCTCTGGACCGCATACTCTGTCCATCTGCGTCCGCCCCAATCTCTTAACCCTCC 2319

Qy        609 CTGCGATCGGCAACTGGGAAGCAATGCCAAGGAGGACAACTGTGGAGTCTGTGCCGGCGA 668
               ||| || |        ||          |     |||   |||||| ||||| || ||||   ||| |
Db       2320 CTGTGACCATGTGGTGAACTCACCTAAGAAGCTGGACAAATGTGGGTGTGTGGGGGCAA 2379

Qy        669 TGGCTCCACCTGCAGG 684
                |||  |  |||| |||
Db       2380 AGGCACTGCCTGTAGG 2395


RESULT 12
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
;  GENERAL INFORMATION:
;     APPLICANT:  DORNER, F.
;     APPLICANT:  SCHEIFLINGER, F.
;     APPLICANT:  FALKNER, F. G.
;     TITLE OF INVENTION:  RECOMBINANT FOWLPOX VIRUS
;     NUMBER OF SEQUENCES:  52
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Foley & Lardner
;       STREET:  1800 Diagonal Road, Suite 500
;       CITY:  Alexandria
;       STATE:  VA
;       COUNTRY:  USA
;       ZIP:  22313-0299
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
        SOFTWARE:  PatentIn Release #1.0, Version #1.25
      CURRENT APPLICATION DATA:
        APPLICATION NUMBER:  US/08/232,463
        FILING DATE:
        CLASSIFICATION:  435
      PRIOR APPLICATION DATA:
        APPLICATION NUMBER:  US/07/935,313
        FILING DATE:
        APPLICATION NUMBER:  EP 91 114 300.6
;       FILING DATE:  26-AUG-1991
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  BENT, Stephen A.
;       REGISTRATION NUMBER:  29,768
;       REFERENCE/DOCKET NUMBER:  30472/114 IMMU
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (703)836-9300
;       TELEFAX:  (703)683-4109
;       TELEX:  899149
;   INFORMATION FOR SEQ ID NO:  14:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  7218 base pairs
;         TYPE:  nucleic acid
;         STRANDEDNESS:  single
;         TOPOLOGY:  linear
;      IMMEDIATE SOURCE:
;         CLONE:  pIZqpt Fls
US-08-232-463-14

  Query Match            1.1%;  Score 56;  DB 1;  Length 7218;
  Best Local Similarity  [illegible]%;  Pred. No. 1e-06;
  Matches  41;  Conservative  215;  Mismatches  190;  Indels  [illegible];  Gaps  [illegible];

Qy       3036 GAGGGAATATCCTGGGATGGA[illegible]CACAGGGAAGCCAATAGTTTGGGAT[illegible] 3095
              [illegible]
Db       1452 GATAGAAGAATTTGGTAC[illegible] 1393

Qy       3096 AATCAGGCAAATGTTGAATAA[illegible]AAAAATGACCTTTATCTGGATGATGA[illegible]CACATTAGTAA 3155
              [illegible]
Db       1392 [illegible] 1333

Qy       3156 [illegible]CAGCCTTTCTTGAGAGCTCTTTAGGCCACTGCAGCAATTCTGCAGGAAGCACCAA[illegible] 3215
              [illegible]
Db       1332 [illegible] 1273

Qy       3216 CTGGGATTTGAAGAATAAGCAGTTTGAAGCAGCATTTAAACAAGGAGCATATAG[illegible]A[illegible]GA 3275
              [illegible]
Db       1272 [illegible] 1213

Qy       3276 TACA[illegible] [illegible]
              [illegible]
Db       1212 [illegible] [illegible]

Qy       3336 CAGCCATGATCTTG[illegible] 3395
              [illegible]
Db       1152 [illegible] 1093

Qy       3396 GCCAACA[illegible] 3455
              [illegible]
Db       1092 [illegible] 1033

Qy       3456 AGCCGAAACAGGGAGT[illegible]AAA 3481
              [illegible]
Db       1032 AATTAATTCTGTGAGG[illegible]AAA 1007


RESULT 13
US-09-491-522-9
; Sequence 9, Application US/09491522
; Patent No. 6428998
;  GENERAL INFORMATION:
;     APPLICANT:  Colige, Alain
;     APPLICANT:  Lapiere, Charles M.
;     APPLICANT:  Prockop, Darwin J.
;     TITLE OF INVENTION:  RECOMBINANT N-PROTEINASE,
;     TITLE OF INVENTION:  AND THE PRODUCTION, METHODS AND USES THEREOF
;     NUMBER OF SEQUENCES:  17
;     CORRESPONDENCE ADDRESS:
;        ADDRESSEE:  Pennie & Edmonds, LLP
;        STREET:  1155 Avenue of the Americas
;        CITY:  New York
;        STATE:  NY
;        COUNTRY:  USA
;        ZIP:  10036-2811
;     COMPUTER READABLE FORM:
;        MEDIUM TYPE:  Diskette
;        COMPUTER:  IBM Compatible
;        OPERATING SYSTEM:  Windows
;        SOFTWARE:  FastSEQ for Windows Version 2.0b
;     CURRENT APPLICATION DATA:
;        APPLICATION NUMBER:  US/09/491,522
;        FILING DATE:
;        CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;        APPLICATION NUMBER:  08/[illegible]
;        FILING DATE:
;     ATTORNEY/AGENT INFORMATION:
;        NAME:  Abrams, Samuel B
;        REGISTRATION NUMBER:  [illegible]
;        REFERENCE/DOCKET NUMBER:  8389-0060-999
;     TELECOMMUNICATION INFORMATION:
;        TELEPHONE:  650-493-4935
;        TELEFAX:  650-493-5556
;        TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  9:
;     SEQUENCE CHARACTERISTICS:
;        LENGTH:  2450 base pairs
;        TYPE:  nucleic acid
;        STRANDEDNESS:  single
;        TOPOLOGY:  linear
US-09-491-522-9

  Query Match            1.1%;  Score 54.[illegible];  DB 4;  Length 24[illegible];
  Best Local Similarity  56.1%;  Pred. No. 1.3e-[illegible];
  Matches  124;  Conservative  [illegible];  Mismatches  [illegible];  Indels  [illegible];  Gaps  [illegible];

Qy        478 [illegible]TAAG 537
              [illegible]
Db       2079 [illegible]ATGAAGCCTATC 2138

Qy        538 [illegible]CATCGCTATCATCGGCATC 594
              [illegible]
Db       2139 [illegible] 2198

Qy        595 [illegible]TGGA 654
              [illegible]
Db       2199 TGCAGGAAGGTGGGCTGTGAC[illegible]GTGATG[illegible]AA[illegible]CA[illegible]GA[illegible]GACAAGTGTGGT 2258

Qy        655 [illegible]
              [illegible]
Db       2259 [illegible]AAGGTGGTCAAGGG 2299


RESULT 14
US-09-491-522-8
; Sequence 8, Application US/09491522
; Patent No. 6428998
;  GENERAL INFORMATION:
;     APPLICANT:  Colige, Alain
;     APPLICANT:  Lapiere, Charles M.
;     APPLICANT:  Prockop, Darwin J.
;     TITLE OF INVENTION:  RECOMBINANT N-PROTEINASE,
;     TITLE OF INVENTION:  AND THE PRODUCTION, METHODS AND USES THEREOF
;     NUMBER OF SEQUENCES:  17
;     CORRESPONDENCE ADDRESS:
;        ADDRESSEE:  Pennie & Edmonds, LLP
;        STREET:  1155 Avenue of the Americas
;        CITY:  New York
;        STATE:  NY
;        COUNTRY:  USA
;        ZIP:  10036-2811
;     COMPUTER READABLE FORM:
;        MEDIUM TYPE:  Diskette
;        COMPUTER:  IBM Compatible
;        OPERATING SYSTEM:  Windows
;        SOFTWARE:  FastSEQ for Windows Version 2.0b
;     CURRENT APPLICATION DATA:
;        APPLICATION NUMBER:  US/09/491,522
;        FILING DATE:
;        CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;        APPLICATION NUMBER:  08/88[illegible],333
;        FILING DATE:
;     ATTORNEY/AGENT INFORMATION:
;        NAME:  Abrams, Samuel B
;        REGISTRATION NUMBER:  30,605
;        REFERENCE/DOCKET NUMBER:  8389-0060-999
;     TELECOMMUNICATION INFORMATION:
```

```
;       TELEPHONE:  650-493-4935
;       TELEFAX:  650-493-5556
;       TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  8:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  4580 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
US-09-491-522-8

  Query Match              1.1%;  Score 54.6;  DB 4;  Length 4580;
  Best Local Similarity   56.1%;  Pred. No. 2e-06;
  Matches  124;  Conservative    0;  Mismatches   94,  Indels    3;  Gaps    1;

Qy     476 CCGTGTGCACTCAAGTGTCATGCACAAGGACAAAACTTGGTGGTGGAGCTGGCACCTAAG 537
           | ||||   ||| | ||| |   |  | |              ||||||     ||   | || |
Db    2079 CGGTGTCATCTCTACTGTGAGTCCAAGGACACCGGGAGGTGGTGTCCATGAAGCGTATG 2138

Qy     538 GTACTGGATGGAACTC---GTTGCAACACGGACTCCTTGGACATGTGTATCAGTGGCATC 594
           |! !  :| || || |    ||| | ||| |||| ||||    | | ||   |  ||||    |
Db    2139 GTGCATGACGGGACACGCTGTTCCTACAAGGACGCCTTCAGCCTCTGCGTGCGTGGGGAC 2198

Qy     595 TGTCAGGCAGTGGGCTGCGATCGGCAACTGGGAAGCAATGCCAAGGAGGACAACTGTGGA 654
           |:   |   |||||:| :|| ||     | ||    ||        ||||||||| |||||
Db    2199 TGCAGGAAGGTGGGCTGTGACGGGGTGATCGGCTCCAGCAAGCAGGAGGACAAGTGTGGT 2258

Qy     655 GTCTGTGCGGGCGATGGCTCCACCTGCAGGCTTGTACGGGG 695
           || |: |  || |:    |:|: |||| | | |!    |||
Db    2259 GTGTGCGGAGGGGACAACTCCACTGCAAGGTGGTCAAGGG 2299


RESULT 15
US-09-369-364A-12
; Sequence 12, Application US/09369364A
; Patent No. 6391610
; GENERAL INFORMATION:
;  APPLICANT:  Apte, Suneel
;  APPLICANT:  Hurskainen, Tiina L.
;  APPLICANT:  Hirohata, Satoshi
;  TITLE OF INVENTION:  Nucleic Acids Encoding Zinc Metalloproteases
;  FILE REFERENCE:  26473/4087/10-30-00
;  CURRENT APPLICATION NUMBER: US/09/369,364A
;  CURRENT FILING DATE:  1999-08-06
;  NUMBER OF SEQ ID NOS:  31
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 12
;   LENGTH: 5804
;   TYPE: DNA
;   ORGANISM: Homo sapiens ADAMTS-9
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (3)..(5648)
;   NAME/KEY: misc_feature
;   LOCATION: (1406)
;   OTHER INFORMATION: n-I
;   NAME/KEY: misc_feature
;   LOCATION: (1563)
;   OTHER INFORMATION: n-I
US-09-369-364A-12

  Query Match              1.0%;  Score 52.4;  DB 4;  Length 5804;
  Best Local Similarity   51.7%;  Pred. No. 1.3e-05;
  Matches  119;  Conservative    0;  Mismatches  111;  Indels    0;  Gaps    0;

Qy     524 AGCTGGCACCTAAGGTACTGGATGGAACTCGTTGCAACACGGACTCCTTGGACATGTGTA 583
           ||||    |    |  ||      ||||||||| !!!    |  |||| !     || || |||
Db    1942 AGCTTCGAGACAGAGTGATAGATGGAACTCCTTGTGGCCAGGACACAAATGATATCTGTG 2001

Qy     584 TCAGTGGCATCTGTCAGGCAGTGGGCTGCGATCGGCAACTGGGAAGCAATGCCAAGGAGG 643
           ||    ||| | || | |   |:  || ||||||        |         || |||  |   |
Db    2002 TCCAGGGCTTTGCCGGCAAGCTGGATGCGATCATGTTTAAACTCAAAAGCCCGGAGAG 2061

Qy     644 ACAACTGTGGAGTCTGTGCCGGCGATGGCTCCACCTGCAGGCTTGTACGGGGACAATCAA 703
           | || || || || ||||  ||||||    ||  : ||||      ||      ||| ||  |
Db    2062 ATAAATGCGGGGTTTGTGGTGGCGATAATTCTTCATGCAAAACAGTGGCAGGAACATTTA 2121

Qy     704 AGTCACACGTTTCTCCTCAAAAAAAGAGAAGAAAATCTAATTGGTGTTCCT 753
           |  ||       || |   |  |  || |   |   ,       |     |||| |  ||
Db    2122 ATACAGTACATTATGGTTACAATACTGTGGTCGAATTCCAGCTGGTGCT 2171

Search completed: March 22, 2003, 10:57:19
Job time : 289 secs
```

```
                    GenCore version 5.1.4_p5_4578
              Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:         March 22, 2003, 09:32:04 ; Search time 274 Seconds
                                           (without alignments)
                                           14366.115 Million cell updates/sec

Title:          US-09-784-358-1
Perfect score:  5076
Sequence:       1 atggcttcct[illegible]..........[illegible] 5076

Scoring table:  IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched:       538826 seqs, 387737923 residues

Total number of hits satisfying chosen parameters:       1077652

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :      Published Applications NA:*
                1:  /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
                2:  /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
                3:  /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
                4:  /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
                5:  /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
                6:  /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
                7:  /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
                8:  /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
                9:  /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
                10: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
                11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
                12: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
                13: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
                14: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                               SUMMARIES
                   %
Result           Query
  No.    Score   Match Length DB  ID                         Description

   1      5076   100.0   5076 10  US-09-784-358-1             Sequence 1, Appli
   2      4766    93.9   4854 10  US-09-784-358-15            Sequence 15, Appl
   3      3755    74.0   8578 10  US-09-784-358-17            Sequence 17, Appl
   4      2490    49.1   2538 10  US-09-784-358-11            Sequence 11, Appl
   5      2180    42.9   2316 10  US-09-784-358-13            Sequence 13, Appl
   6      2119    41.7   2175 10  US-09-784-358-7             Sequence 7, Appli
   7      1809    35.6   1853 10  US-09-784-358-9             Sequence 9, Appli
   8      1327    26.1   1341 10  US-09-784-358-3             Sequence 3, Appli
   9    1020.2    20.1   5289  9  US-10-044-807-1            Sequence 1, Appli
  10      1017    20.0   1119 10  US-09-784-358-5             Sequence 5, Appli
  11     925.4    18.2   5769 10  US-09-764-853-134           Sequence 134, App
  12     924.2    18.2   5720 10  US-09-800-729-18            Sequence 18, Appl
  13       913    18.0   2080  9  US-09-796-753-17           Sequence 17, Appl
  14       912    18.0   2080  9  US-09-796-753-19           Sequence 19, Appl
  15     843.2    16.6   3594 10  US-09-800-729-53            Sequence 53, Appl
  16       738    14.5   1869  9  US-09-978-295A-300         Sequence 300, App
  17       738    14.5   1869  9  US-09-978-667-300          Sequence 300, App
  18       738    14.5   1869  9  US-09-978-192A-300         Sequence 300, App
  19       738    14.5   1869  9  US-09-999-832A-300         Sequence 300, App
  20       738    14.5   1869  9  US-09-978-189-300          Sequence 300, App
  21       738    14.5   1869  9  US-10-174-[illegible]-435  Sequence 435, App
  22       738    14.5   1869  9  US-10-176-558-435          Sequence 435, App
  23       738    14.5   1869  9  US-10-175-[illegible]-435  Sequence 435, App
  24       738    14.5   1869  9  US-10-173-206-435          Sequence 435, App
  25       738    14.5   1869  9  US-10-175-738-435          Sequence 435, App
  26       738    14.5   1869  9  US-10-175-752-435          Sequence 435, App
  27       738    14.5   1869  9  US-10-176-482-435          Sequence 435, App
  28       738    14.5   1869  9  US-10-176-757-435          Sequence 435, App
  29       738    14.5   1869  9  US-10-176-913-435          Sequence 435, App
  30       738    14.5   1869  9  US-10-180-552-435          Sequence 435, App
  31       738    14.5   1869  9  US-10-180-557-435          Sequence 435, App
  32       738    14.5   1869  9  US-10-174-700-435          Sequence 435, App
  33       738    14.5   1869  9  US-10-174-572-435          Sequence 435, App
  34       738    14.5   1869  9  US-10-174-579-435          Sequence 435, App
  35       738    14.5   1869  9  US-10-174-582-435          Sequence 435, App
  36       738    14.5   1869  9  US-10-174-588-435          Sequence 435, App
  37       738    14.5   1869  9  US-10-175-739-435          Sequence 435, App
  38       738    14.5   1869  9  US-10-175-740-435          Sequence 435, App
  39       738    14.5   1869  9  US-10-175-743-435          Sequence 435, App
  40       738    14.5   1869  9  US-10-176-488-435          Sequence 435, App
  41       738    14.5   1869  9  US-10-176-492-435          Sequence 435, App
  42       738    14.5   1869  9  US-10-176-747-435          Sequence 435, App
  43       738    14.5   1869  9  US-10-176-750-435          Sequence 435, App
  44       738    14.5   1869  9  US-10-176-985-435          Sequence 435, App
  45       738    14.5   1869  9  US-10-176-987-435          Sequence 435, App

                               ALIGNMENTS


RESULT 1
US-09-784-358-1
; Sequence 1, Application US/09784358
; Patent No. US2002[illegible]A1
; GENERAL INFORMATION:
;  APPLICANT: Donoho, Gregory
;  APPLICANT: Scoville, John
;  APPLICANT: Turner, C. Alexander Jr.
;  APPLICANT: Friedrich, Glenn
;  APPLICANT: Zambrowicz, Brian
;  APPLICANT: Sands, Arthur T.
;  TITLE OF INVENTION: NOVEL HUMAN THROMBOSPONDIN REPEAT PROTEINS AND
;  TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING THE SAME
;  FILE REFERENCE: LEX-0134-USA
;  CURRENT APPLICATION NUMBER: US/09/784,358
;  CURRENT FILING DATE: 2001-02-15
;  PRIOR APPLICATION NUMBER: US 60/183,282
;  PRIOR FILING DATE: 2000-02-17
;  NUMBER OF SEQ ID NOS: 17
;  SOFTWARE: FastSEQ for Windows Version 4.
; SEQ ID NO 1
;   LENGTH: 5076
;   TYPE: DNA
;   ORGANISM: homo sapiens
US-09-784-358-1

  Query Match             100.0%;  Score 5076;  DB 10;  Length 5076;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 5076;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGGCTTCCTGGACGAG[illegible]GATAGGGA[illegible] 60
Db          1 ATGGCTTCCTGGACGAGCC[illegible]GATAGGGA[illegible] 60

Qy         61 [illegible]GCAGACCACAGC[illegible]AGAAA[illegible]GCATT 120
Db         61 [illegible]GCAGACCACAGCTCAGAAA[illegible]GCATT 120

Qy        121 [illegible]CAGGAAGTTT[illegible]GAAGACACAA[illegible]TAT 180
Db        121 [illegible]CAGGAAGTTT[illegible]GAAGACACAA[illegible]TAT 180
```

```
Qy        181 GATGACCAGACCTCAAGAAACACTCGTTCAGATGAAGACAAAGATGGCAACTGGGATGCT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GATGACCAGACCTCAAGAAACACTCGTTCAGATGAAGACAAAGATGGCAACTGGGATGCT 240

Qy        241 TGGGGCGACTGGAGTGACTGCTCCCGGACCTGTGGGGGAGGAGCATCATATTCTCTGCGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGGGGCGACTGGAGTGACTGCTCCCGGACCTGTGGGGGAGGAGCATCATATTCTCTGCGG 300

Qy        301 AGATGTTTGACTGGAAGGAATTGTGAAGGGCAGAACATTCGGTACAAGACATGCAGCAAT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGATGTTTGACTGGAAGGAATTGTGAAGGGCAGAACATTCGGTACAAGACATGCAGCAAT 360

Qy        361 CATGACTGCCCTCCAGATGCAGAAGATTTCAGAGCCCAGCAGTGCTCAGCCTACAATGAT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CATGACTGCCCTCCAGATGCAGAAGATTTCAGAGCCCAGCAGTGCTCAGCCTACAATGAT 420

Qy        421 GTCCAGTATCAGGGGCATTACTATGAATGGCTTCCACGATATAATGATCCTGCTGCCCCG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GTCCAGTATCAGGGGCATTACTATGAATGGCTTCCACGATATAATGATCCTGCTGCCCCG 480

Qy        481 TGTGCACTCAAGTGTCATGCACAAGGACAAAACTTGGTGGTGGAGCTGGCACCTAAGGTA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TGTGCACTCAAGTGTCATGCACAAGGACAAAACTTGGTGGTGGAGCTGGCACCTAAGGTA 540

Qy        541 CTGGATGGAACTCGTTGCAACACGGACTCCTTGGACATGTGTATCAGTGGCATCTGTCAG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTGGATGGAACTCGTTGCAACACGGACTCCTTGGACATGTGTATCAGTGGCATCTGTCAG 600

Qy        601 GCAGTGGGCTGCGATCGGCAACTGGGAAGCAATGCCAAGGAGGACAACTGTGGAGTCTGT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GCAGTGGGCTGCGATCGGCAACTGGGAAGCAATGCCAAGGAGGACAACTGTGGAGTCTGT 660

Qy        661 GCCGGCGATGGCTCCACCTGCAGGCTTGTACGGGGACAATCAAAGTCACACGTTTCTCCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GCCGGCGATGGCTCCACCTGCAGGCTTGTACGGGGACAATCAAAGTCACACGTTTCTCCT 720

Qy        721 GAAAAAAGAGAAGAAAATGTAATTGCTGTTCCTTTGGGAAGTCGAAGTGTGAGAATTACA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GAAAAAAGAGAAGAAAATGTAATTGCTGTTCCTTTGGGAAGTCGAAGTGTGAGAATTACA 780

Qy        781 GTGAAAGGACCTGCCCACCTCTTTATTGAATCAAAAACACTTCAAGGAAGCAAAGGAGAA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GTGAAAGGACCTGCCCACCTCTTTATTGAATCAAAAACACTTCAAGGAAGCAAAGGAGAA 840

Qy        841 CACAGCTTTAACAGCCCCGGCGTCTTTGTCGTAGAAAACACAACAGTGGAATTTCAGAGG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CACAGCTTTAACAGCCCCGGCGTCTTTGTCGTAGAAAACACAACAGTGGAATTTCAGAGG 900

Qy        901 GGCTCCGAGAGGCAAACTTTTAAGATTCCAGGACCTCTGATGGCTGATTTCATCTTCAAG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GGCTCCGAGAGGCAAACTTTTAAGATTCCAGGACCTCTGATGGCTGATTTCATCTTCAAG 960

Qy        961 ACCAGGTACACTGCAGCCAAAGACAGCGTGGTTCAGTTCTTCTTTACCAGCCCATCAGT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 ACCAGGTACACTGCAGCCAAAGACAGCGTGGTTCAGTTCTTCTTTACCAGCCCATCAGT 1020

Qy       1021 CATCAGTGGAGACAAACTGACTTCTTTCCCTGCACTGTGACGTGTGGAGGAGGTTATCAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CATCAGTGGAGACAAACTGACTTCTTTCCCTGCACTGTGACGTGTGGAGGAGGTTATCAG 1080

Qy       1081 CTCAATTCTGCTGAATGTGTGGATATCCGCTTGAAGAGGGTAGTTCCTGACCATTATTGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CTCAATTCTGCTGAATGTGTGGATATCCGCTTGAAGAGGGTAGTTCCTGACCATTATTGT 1140

Qy       1141 CACTACTACCCTGAAAATGTAAAACCAAAACCAAAACTGAAGGAATGCAGCATGGATCCC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CACTACTACCCTGAAAATGTAAAACCAAAACCAAAACTGAAGGAATGCAGCATGGATCCC 1200

Qy       1201 TGCCCATCAAGTGATGGATTTAAAGAGATAATGCCCTATGACCACTTCCAACCTCTTCCT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TGCCCATCAAGTGATGGATTTAAAGAGATAATGCCCTATGACCACTTCCAACCTCTTCCT 1260

Qy       1261 CGCTGGGAACATAATCCTTGGACTGCATGTTCCGTGTCCTGTGGAGGAGGGATTCAGAGA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CGCTGGGAACATAATCCTTGGACTGCATGTTCCGTGTCCTGTGGAGGAGGGATTCAGAGA 1320

Qy       1321 CGGAGCTTTGTGTGTGTAGAGGAATCCATGCATGGAGAGATATTGCAGGTGGAAGAATGG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CGGAGCTTTGTGTGTGTAGAGGAATCCATGCATGGAGAGATATTGCAGGTGGAAGAATGG 1380

Qy       1381 AAGTGCATGTACGCACCCAAACCCAAGGTTATGCAAACTTGTAATCTGTTTGATTGCCCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 AAGTGCATGTACGCACCCAAACCCAAGGTTATGCAAACTTGTAATCTGTTTGATTGCCCC 1440

Qy       1441 AAGTGGATTGCCATGGAGTGGTCTCAGTGCACAGTGACTTGTGGCCGAGGGTTACGGTAC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 AAGTGGATTGCCATGGAGTGGTCTCAGTGCACAGTGACTTGTGGCCGAGGGTTACGGTAC 1500

Qy       1501 CGGGTTGTTCTGTGTATTAACCACCGCGGAGAGCATGTTGGGGGCTGCAATCCACAACTG 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 CGGGTTGTTCTGTGTATTAACCACCGCGGAGAGCATGTTGGGGGCTGCAATCCACAACTG 1560

Qy       1561 AAGTTACACATCAAAGAAGAATGTGTCATTCCCATCCCGTGTTATAAACCAAAAGAAAAA 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 AAGTTACACATCAAAGAAGAATGTGTCATTCCCATCCCGTGTTATAAACCAAAAGAAAAA 1620

Qy       1621 AGTCCAGTGGAAGCAAAATTGCCTTGGCTGAAACAAGCACAAGAACTAGAAGAGACCAGA 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 AGTCCAGTGGAAGCAAAATTGCCTTGGCTGAAACAAGCACAAGAACTAGAAGAGACCAGA 1680

Qy       1681 ATAGCAACAGAAGAACCAACGTTCATTCCAGAACCCTGGTCAGCCTGCAGTACCACGTGT 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 ATAGCAACAGAAGAACCAACGTTCATTCCAGAACCCTGGTCAGCCTGCAGTACCACGTGT 1740

Qy       1741 GGGCCAGGTGTGCAGGTCCGCGAGGTGAAGTGCCGTGTGCTCCTCACATTCACGCAGACT 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 GGGCCAGGTGTGCAGGTCCGCGAGGTGAAGTGCCGTGTGCTCCTCACATTCACGCAGACT 1800

Qy       1801 GAGACTGAGCTGCCCGAGGAAGAGTGTGAAGGCCCCAAGCTGCCCACCGAACGGCCCTGC 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 GAGACTGAGCTGCCCGAGGAAGAGTGTGAAGGCCCCAAGCTGCCCACCGAACGGCCCTGC 1860

Qy       1861 CTCCTGGAAGCATGTGATGAGAGCCCGGCCTCCCGAGAGCTAGACATCCCTCTCCCTGAG 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 CTCCTGGAAGCATGTGATGAGAGCCCGGCCTCCCGAGAGCTAGACATCCCTCTCCCTGAG 1920

Qy       1921 GACAGTGAGACGACTTACGACTGGGAGTACGCTGGGTTCACCCCTTGCACAGCAACATGC 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 GACAGTGAGACGACTTACGACTGGGAGTACGCTGGGTTCACCCCTTGCACAGCAACATGC 1980

Qy       1981 TTGGGAGGCCATCAAGAAGCCATAGCAGTGTGCTTACATATCCAGACCCAGCAGACAGTC 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 TTGGGAGGCCATCAAGAAGCCATAGCAGTGTGCTTACATATCCAGACCCAGCAGACAGTC 2040

Qy       2041 AATGACAGCTTGTGTGATATGGTCCACCGTCCTCCAGCCATGAGCCAGGCCTGTAACACA 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 AATGACAGCTTGTGTGATATGGTCCACCGTCCTCCAGCCATGAGCCAGGCCTGTAACACA 2100

Qy       2101 GAGCCCTGTCCCCCCAGGTGGCATGTGGCCTCTTGGGGGCCCTGCTCAGCTACCTGTGGA 2160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 GAGCCCTGTCCCCCCAGGTGGCATGTGGCCTCTTGGGGGCCCTGCTCAGCTACCTGTGGA 2160

Qy       2161 GTTGGAATTCAGACCCGAGATGTGTACTGCCTGCACCCAGGGGAGACCCCTGCCCCTCCT 2220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 GTTGGAATTCAGACCCGAGATGTGTACTGCCTGCACCCAGGGGAGACCCCTGCCCCTCCT 2220

Qy       2221 CAGCACTGCCGACATGAAAAGCCCCATGCTTTACAAGCATGCAATCAGTTTCACTGCCCT 2280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 CAGCACTGCCGACATGAAAAGCCCCATGCTTTACAAGCATGCAATCAGTTTCACTGCCCT 2280

Qy       2281 CCTGCCTGCCACATTGAAGAATGGCAGCAGTGTTCCAGGACTTGTGGCGGGGGAACTCAG 2340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 CCTGCCTGCCACATTGAAGAATGGCAGCAGTGTTCCAGGACTTGTGGCGGGGGAACTCAG 2340

Qy       2341 AACAGAAGAGTCACCTGTCGGCAGCTGCTAACGGATGGCAGCTTTTTGAATCTCTCAGAT 2400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 AACAGAAGAGTCACCTGTCGGCAGCTGCTAACGGATGGCAGCTTTTGAATCTCTCAGAT 2400

Qy       2401 GAATTGTGCCAAGGACCTAAGCCATCCTCTCACAAGTCCTGTGCCAGGACAGACTGTCCT 2460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 GAATTGTGCCAAGGACCTAAGCCATCCTCTCACAAGTCCTGTGCCAGGACAGACTGTCCT 2460

Qy       2461 CCACATTAGCTGTGGGAGACTGGTCGAAGTGTTCTGTCAGTTGTGGTGTTGGAATCCAG 2520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 CCACATTAGCTGTGGGAGACTGGTCGAAGTGTTCTGTCAGTTGTGGTGTTGGAATCCAG 2520

Qy       2521 AGAAGAAAGCAGGTGTGTCAAAGGCTGGCAGCCAAAGGTCGGCGCATCCCCCTCAGTGAG 2580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 AGAAGAAAGCAGGTGTGTCAAAGGCTGGCAGCCAAAGGTCGGCGCATCCCCCTCAGTGAG 2580

Qy       2581 ATGATGTGCAGGGATCTACCAGGGCTCCCTCTTGTAAGATCTTGCCAGATCCCTGACCGC 2640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 ATGATGTGCAGGGATCTACCAGGGCTCCCTCTTGTAAGATCTTGCCAGATCCCTGACCGC 2640

Qy       2641 AGTAAAATCAAATCAGAGATGAAGACAAAACTTGGTGAGCAGGGTCCGCAGATCCTCAGT 2700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 AGTAAAATCAAATCAGAGATGAAGACAAAACTTGGTGAGCAGGGTCCGCAGATCCTCAGT 2700

Qy       2701 GTCCAGAGAGTCTACATTCAGACAAGGGAAGACAAGCGTATTAACCTGACCATTGGTAGC 2760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 GTCCAGAGAGTCTACATTCAGACAAGGGAAGAGAAGCGTATTAACCTGACCATTGGTAGC 2760

Qy       2761 AGAGCCTATTTGCTGCCTAACACATCCGTGATTATTAAGTGCCCCGTGCGACGATTCCAG 2820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 AGAGCCTATTTGCTGCCTAACACATCCGTGATTATTAAGTGCCCCGTGCGACGATTCCAG 2820

Qy       2821 AAATCTCTGATCCAGTGGGAGAAGGATGGCCGTTGCCTGCAGAACTCCAAACGGCTTGGC 2880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 AAATCTCTGATCCAGTGGGAGAAGGATGGCCGTTGCCTGCAGAACTCCAAACGGCTTGGC 2880

Qy       2881 ATCACCAAGTCAGGCTTACTAAAAATCCACGGTCTTGACTGCCCCCGACATCCCAGGTAC 2940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 ATCACCAAGTCAGGCTTACTAAAAATCCACGGTCTTGACTGCCCCCGACATCCCAGGTAC 2940

Qy       2941 CGGTGCATTGCAGGCTCTGCACAGGAAACAGTTGTGCTCAAGCTCATTGGTACTGACAAC 3000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 CGGTGCATTGCAGGCTCTGCACAGGAAACAGTTGTGCTCAAGCTCATTGGTACTGACAAC 3000

Qy       3001 CGGCTCATCGCAGCCCAGCCTCAGGGAGCCTATGAGGGAATATCCTGGGATGGACCCAC 3060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 CGGCTCATCGCAGCCCAGCCTCAGGGAGCCTATGAGGGAATATCCTGGGATGGACCCAC 3060

Qy       3061 AGCGAAGCCAATAGTTTGGGAGTCACATGGCACAAAATGAGGCAAATGTGGAATAACAAA 3120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 AGCGAAGCCAATAGTTTGGGAGTCACATGGCACAAAATGAGGCAAATGTGGAATAACAAA 3120

Qy       3121 AATGACCTTTATCTGGATGATGACCACATTAGTAACCAGCCTTCTTGAGAGCTCCTGTTA 3180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3121 AATGACCTTTATCTGGATGATGACCACATTAGTAACCAGCCTTCTTGAGAGCTCCTGTTA 3180

Qy       3181 GGCCACTGGAGCAATTCCTCAGGAAGCACCAACTCCTGGGAGTTGAAGAATAAGCAGTTT 3240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3181 GGCCACTGGAGCAATTCCTCAGGAAGCACCAACTCCTGGGAGTTGAAGAATAAGCAGTTT 3240

Qy       3241 GAAGCAGCAGTTAAACAAGGAGCATATAGCATGGATACAGCCCAGTTTGATGAGCTGATA 3300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3241 GAAGCAGCAGTTAAACAAGGAGCATATAGCATGGATACAGCCCAGTTTGATGAGCTGATA 3300

Qy       3301 AGAAACATGAGTCAGCTCATGGAAACCGGAGAGGTCAGCGATGATCTTGCGTCCCAGCTG 3360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3301 AGAAACATGAGTCAGCTCATGGAAACCGGAGAGGTCAGCGATGATCTTGCGTCCCAGCTG 3360

Qy       3361 ATATATCAGCTGGTGGCCGAATTAGCCAAGGCACAGCCAACACACATGCAGTGGCGGGGC 3420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3361 ATATATCAGCTGGTGGCCGAATTAGCCAAGGCACAGCCAACACACATGCAGTGGCGGGGC 3420

Qy       3421 ATCCAGGAAGAGACACCTCCTGCTGCTCAGCTCAGAGGGAAACAGGGAGTGTGTGTCCCAA 3480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3421 ATCCAGGAAGAGACACCTCCTGCTGCTCAGCTCAGAGGGAAACAGGGAGTGTGTGTCCCAA 3480

Qy       3481 AGCTGGCATGCAAAAAACTCAGGCAAGCTGACATTCAAGCCGAAAGGACCTGTTCTCATG 3540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3481 AGCTGGCATGCAAAAAACTCAGGCAAGCTGACATTCAAGCCGAAAGGACCTGTTCTCATG 3540

Qy       3541 AGGCAAAGCCAACCTCCCTCAATTTCATTTAATAAAACAATAAATCCCAGGATTGGAAAT 3600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3541 AGGCAAAGCCAACCTCCCTCAATTTCATTTAATAAAACAATAAATCCCAGGATTGGAAAT 3600

Qy       3601 ACAGTATACATTACAAAAAGGACAGAGGTCATCAATATACTGGTGACCTTATTACCCCCC 3660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3601 ACAGTATACATTACAAAAAGGACAGAGGTCATCAATATACTGGTGACCTTATTACCCCCC 3660

Qy       3661 AGTGAGGCCACATATACATGGACCAAGGATGGAACCTTGTTACAGCCCTCAGTAAAAATA 3720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3661 AGTGAGGCCACATATACATGGACCAAGGATGGAACCTTGTTACAGCCCTCAGTAAAAATA 3720

Qy       3721 ATTTTGGATGGAACTGGGAAGATACAGATACAGAATCCTACAAGGAAAGAACAAGGCATA 3780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3721 ATTTTGGATGGAACTGGGAAGATACAGATACAGAATCCTACAAGGAAAGAACAAGGCATA 3780

Qy       3781 TATGAATGTTCTGTAGCTAATCATCTTGGTTCAGATGTGGAAAGTTCTTCTGTGCTGTAT 3840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3781 TATGAATGTTCTGTAGCTAATCATCTTGGTTCAGATGTGGAAAGTTCTTCTGTGCTGTAT 3840

Qy       3841 GCAGAGGCACCTGTCATCTTGTCTGTTGAAAGAAATATCACCAAACCAGAGCACAACCAT 3900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3841 GCAGAGGCACCTGTCATCTTGTCTGTTGAAAGAAATATCACCAAACCAGAGCACAACCAT 3900

Qy       3901 CTGTCTGTTGTGGTTGGAGGCATCGTGGAGGCAGCCCTTGGAGCAAACGTGACAATCCGA 3960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3901 CTGTCTGTTGTGGTTGGAGGCATCGTGGAGGCAGCCCTTGGAGCAAACGTGACAATCCGA 3960

Qy       3961 TGTCCTGTAAAAGTGTGCCTCAGCCTAATATAACTTGGCTTGAAGAGAGGAGGATCTCTG 4020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3961 TGTCCTGTAAAAGTGTGCCTCAGCCTAATATAACTTGGCTTGAAGAGAGGAGGATCTCTG 4020

Qy       4021 AGTGGCAATGTTTCCTTGCTTTTCAATGGATCCCTGTTGTTGCAGAATGTTTCCCTTGAA 4080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4021 AGTGGCAATGTTTCCTTGCTTTTCAATGGATCCCTGTTGTTGCAGAATGTTTCCCTTGAA 4080

Qy       4081 AATGAAGGAACCTACGTCTGCATAGCCACCAATGCTCTTGGAAAGGCAGTGGCAACATCT 4140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4081 AATGAAGGAACCTACGTCTGCATAGCCACCAATGCTCTTGGAAAGGCAGTGGCAACATCT 4140

Qy       4141 GTACTCCACTTGCTGGAACGAAGATGGCCAGAGAGTAGAATCGTATTTCTGCAAGGACAT 4200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4141 GTACTCCACTTGCTGGAACGAAGATGGCCAGAGAGTAGAATCGTATTTCTGCAAGGACAT 4200

Qy       4201 AAAAAGTACATTCTCCAGGCAACCAACACTAGAACCAACAGCAATGACCCAACAGGAGAA 4260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4201 AAAAAGTACATTCTCCAGGCAACCAACACTAGAACCAACAGCAATGACCCAACAGGAGAA 4260

Qy       4261 CCCCCGCCTCAAGAGCCTTTTGGGAGCCTGGTAACTGGTTCACATTCTTCTGCCACCTGT 4320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4261 CCCCCGCCTCAAGAGCCTTTTGGGAGCCTGGTAACTGGTTCACATTCTTCTGCCACCTGT 4320

Qy       4321 GGTCATTTGGGAGCCCGCATTCAGAGACCCCAGTGTGTGATGGCCAATGGGCAGGAAGTG 4380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4321 GGTCATTTGGGAGCCCGCATTCAGAGACCCCAGTGTGTGATGGCCAATGGGCAGGAAGTG 4380

Qy       4381 AGTGAGGCCCTGTGTGATCACCTCCAGAAGCCACTGGCTGGTTTGAGCCCCTGTAACATC 4440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4381 AGTGAGGCCCTGTGTGATCACCTCCAGAAGCCACTGGCTGGTTTGAGCCCCTGTAACATC 4440

Qy       4441 CGGGACTGCCCAGCGAGGCGGTTCACAAGTGTGTGGTCACAGTGACTCTGTGGCTTGCGGT 4500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4441 CGGGACTGCCCAGCGAGGCGGTTCACAAGTGTGTGGTCACAGTGACTCTGTGGCTTGCGGT 4500

Qy       4501 GAAGGATACCACAGTCGGCAGGTGACGTGCAAGCGGACAAAAGCCAATGGAACTGTGCAG 4560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4501 GAAGGATACCACAGTCGGCAGGTGACGTGCAAGCGGACAAAAGCCAATGGAACTGTGCAG 4560
```

```
Qy       4561 GTGGTGTCTCCAAGAGCATGTGCCCCTAAAGACCGGCCTCTGGGAAGAAAACCATGTTTT 4620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4561 GTGGTGTCTCCAAGAGCATGTGCCCCTAAAGACCGGCCTCTGGGAAGAAAACCATGTTTT 4620

Qy       4621 GGTCATCCATGTGTTCAGTGGGAACCAGGGAACCGGTGTCCTGGACGTTGCATGGGCCGT 4680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4621 GGTCATCCATGTGTTCAGTGGGAACCAGGGAACCGGTGTCCTGGACGTTGCATGGGCCGT 4680

Qy       4681 GCTGTGAGGATGCAGCAGCGTCACACAGCTTGTCAACACAACAGCTCTGACTCCAACTGT 4740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4681 GCTGTGAGGATGCAGCAGCGTCACACAGCTTGTCAACACAACAGCTCTGACTCCAACTGT 4740

Qy       4741 GATGACAGAAAGAGACCCACCTTAAGAAGGAACTGCACATCAGGGGCCTGTGATGTGTGT 4800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4741 GATGACAGAAAGAGACCCACCTTAAGAAGGAACTGCACATCAGGGGCCTGTGATGTGTGT 4800

Qy       4801 TGGCACACAGGCCCTTGGAAGCCCTGTACAGCAGCCTGTGGCAGGGGTTTCCAGTCTCGG 4860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4801 TGGCACACAGGCCCTTGGAAGCCCTGTACAGCAGCCTGTGGCAGGGGTTTCCAGTCTCGG 4860

Qy       4861 AAAGTCGACTGTATCCACACAAGGAGTTGCAAACCTGTGGCCAAGAGACACTGTGTACAG 4920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4861 AAAGTCGACTGTATCCACACAAGGAGTTGCAAACCTGTGGCCAAGAGACACTGTGTACAG 4920

Qy       4921 AAAAAGAAACCAATTTCCTGGCGGCACTGTCTTGGGCCCTCCTGTGATAGAGACTGCACA 4980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4921 AAAAAGAAACCAATTTCCTGGCGGCACTGTCTTGGGCCCTCCTGTGATAGAGACTGCACA 4980

Qy       4981 GACACAACTCACTACTGTATGTTTGTAAAACATCTTAATTTGTGTTCTCTAGACCGCTAC 5040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4981 GACACAACTCACTACTGTATGTTTGTAAAACATCTTAATTTGTGTTCTCTAGACCGCTAC 5040

Qy       5041 AAACAAAGGTGCTGCCAGTCATGTCAAGAGGGATAA 5076
              ||||||||||||||||||||||||||||||||||||
Db       5041 AAACAAAGGTGCTGCCAGTCATGTCAAGAGGGATAA 5076


RESULT 2
US-09-784-358-15
; Sequence 15, Application US/09784358
; Patent No. US20020099027A1
; GENERAL INFORMATION:
;  APPLICANT: Donoho, Gregory
;  APPLICANT:  Scoville, John
;  APPLICANT:  Turner, C. Alexander Jr.
;  APPLICANT:  Friedrich, Glenn
;  APPLICANT:  Zambrowicz, Brian
;  APPLICANT:  Sands, Arthur T.
;  TITLE OF INVENTION: NOVEL HUMAN THROMBOSPONDIN REPEAT PROTEINS AND
;  TITLE OF INVENTION:  POLYNUCLEOTIDES ENCODING THE SAME
;  FILE REFERENCE: LEX-0134-USA
;  CURRENT APPLICATION NUMBER: US/09/784,358
;  CURRENT FILING DATE:  2001-02-15
;  PRIOR APPLICATION NUMBER: US 60/183,282
;  PRIOR FILING DATE: 2000-02-17
;  NUMBER OF SEQ ID NOS: 17
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15
;   LENGTH: 4854
;   TYPE: DNA
;   ORGANISM: homo sapiens
US-09-784-358-15

  Query Match             93.9%;  Score 4766;  DB 10;  Length 4854;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 4766;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        311 CTGGAAGGAATTGTGAAGGGCAGAACATTCGGTACAAGACATGCAGCAATCATGACTGCC 370
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         89 CTGGAAGGAATTGTGAAGGGCAGAACATTCGGTACAAGACATGCAGCAATCATGACTGCC 148

Qy        371 CTCCAGATGCAGAAGATTTCAGAGCCCAGCAGTGCTCAGCCTACAATGATGTCCAGTATC 430
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        149 CTCCAGATGCAGAAGATTTCAGAGCCCAGCAGTGCTCAGCCTACAATGATGTCCAGTATC 208

Qy        431 AGGGGCATTACTATGAATGGCTTCCACGATATAATGATCCTGCTGCCCCGTGTGCACTCA 490
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        209 AGGGGCATTACTATGAATGGCTTCCACGATATAATGATCCTGCTGCCCCGTGTGCACTCA 268

Qy        491 AGTGTCATGCACAAGGACAAAACTTGGTGGTGGAGCTGGCACCTAAGGTACTGGATGGAA 550
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        269 AGTGTCATGCACAAGGACAAAACTTGGTGGTGGAGCTGGCACCTAAGGTACTGGATGGAA 328

Qy        551 CTCGTTGCAACACGGACTCCTTGGACATGTGTATCAGTGGCATCTGTCAGGCAGTGGGCT 610
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        329 CTCGTTGCAACACGGACTCCTTGGACATGTGTATCAGTGGCATCTGTCAGGCAGTGGGCT 388

Qy        611 GCGATCGGCAACTGGGAAGCAATGCCAAGGAGGACAACTGTGGAGTCTGTGCCGGCGATG 670
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        389 GCGATCGGCAACTGGGAAGCAATGCCAAGGAGGACAACTGTGGAGTCTGTGCCGGCGATG 448

Qy        671 GCTCCACCTGCAGGCTTGTACGGGGACAATCAAAGTCACACGTTTCTCCTGAAAAAAGAG 730
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        449 GCTCCACCTGCAGGCTTGTACGGGGACAATCAAAGTCACACGTTTCTCCTGAAAAAAGAG 508

Qy        731 AAGAAAATGTAATTGCTGTTCCTTTGGGAAGTCGAAGTGTGAGAATTACAGTGAAAGGAC 790
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        509 AAGAAAATGTAATTGCTGTTCCTTTGGGAAGTCGAAGTGTGAGAATTACAGTGAAAGGAC 568

Qy        791 CTGCCCACCTCTTTATTGAATCAAAAACACTTCAAGGAAGCAAAGGAGAACACAGCTTTA 850
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        569 CTGCCCACCTCTTTATTGAATCAAAAACACTTCAAGGAAGCAAAGGAGAACACAGCTTTA 628

Qy        851 ACAGCCCCGGCGTCTTTGTCGTAGAAAACACAACAGTGGAATTTCAGAGGGGCTCCGAGA 910
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        629 ACAGCCCCGGCGTCTTTGTCGTAGAAAACACAACAGTGGAATTTCAGAGGGGCTCCGAGA 688

Qy        911 GGCAAACTTTTAAGATTCCAGGACCTCTGATGGCTGATTTCATCTTCAAGACCAGGTACA 970
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        689 GGCAAACTTTTAAGATTCCAGGACCTCTGATGGCTGATTTCATCTTCAAGACCAGGTACA 748

Qy        971 CTGCAGCCAAAGACAGCGTGGTTCAGTTCTTCTTTTACCAGCCCATCAGTCATCAGTGGA 1030
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        749 CTGCAGCCAAAGACAGCGTGGTTCAGTTCTTCTTTTACCAGCCCATCAGTCATCAGTGGA 808

Qy       1031 GACAAACTGACTTCTTTCCCTGCACTGTGACGTGTGGAGGAGGTTATCAGCTCAATTCTG 1090
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        809 GACAAACTGACTTCTTTCCCTGCACTGTGACGTGTGGAGGAGGTTATCAGCTCAATTCTG 868

Qy       1091 CTGAATGTGTGGATATCCGCTTGAAGAGGGTAGTTCCTGACCATTATTGTCACTACTACC 1150
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        869 CTGAATGTGTGGATATCCGCTTGAAGAGGGTAGTTCCTGACCATTATTGTCACTACTACC 928

Qy       1151 CTGAAAATGTAAAACCAAAACCAAAACTGAAGGAATGCAGCATGGATCCCTGCCCATCAA 1210
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        929 CTGAAAATGTAAAACCAAAACCAAAACTGAAGGAATGCAGCATGGATCCCTGCCCATCAA 988

Qy       1211 GTGATGGATTTAAAGAGATAATGCCCTATGACCACTTCCAACCTCTTCCTCGCTGGGAAC 1270
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        989 GTGATGGATTTAAAGAGATAATGCCCTATGACCACTTCCAACCTCTTCCTCGCTGGGAAC 1048

Qy       1271 ATAATCCTTGGACTGCATGTTCCGTGTCCTGTGGAGGAGGGATTCAGAGACGGAGCTTTG 1330
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1049 ATAATCCTTGGACTGCATGTTCCGTGTCCTGTGGAGGAGGGATTCAGAGACGGAGCTTTG 1108

Qy       1331 TGTGTGTAGAGGAATCCATGCATGGAGAGATATTGCAGGTGGAAGAATGGAAGTGCATGT 1390
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1109 TGTGTGTAGAGGAATCCATGCATGGAGAGATATTGCAGGTGGAAGAATGGAAGTGCATGT 1168

Qy       1391 ACGCACCCAAACCCAAGGTTATGCAAACTTGTAATCTGTTTGATTGCCCCAAGTGGATTG 1450
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1169 ACGCACCCAAACCCAAGGTTATGCAAACTTGTAATCTGTTTGATTGCCCCAAGTGGATTG 1228

Qy       1451 CCATGGAGTGGTCTCAGTGCACAGTGACTTGTGGCCGAGGGTTACGGTACCGGGTTGTTC 1510
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1229 CCATGGAGTGGTCTCAGTGCACAGTGACTTGTGGCCGAGGGTTACGGTACCGGGTTGTTC 1288
```

```
Qy       1511 TGTGTATTAACCAGCAGGTAGAGCATGTTGGGGCTGCAATCCACAACTGAAGTTACATA 1570
Db       1289 TGTGTATTAACCACCCGGTACACCATGTTGGGGCTGCAATCCACAACTGAAGTTACATA 1348

Qy       1571 TCAAAGAAGAATGTGTCATCCCATCCCGTGTTATAAACCAAAAGAAAAAAGCCCAGTGG 1630
Db       1349 TCAAAGAAGAATGTGTCATTCCCATCCCGTGTTATAAACCAAAAGAAAAAAGTCCAGTGG 1408

Qy       1631 AAGCAAAATTGCCTTGGCTGAAACAAGCACAAGAACTAGAAGAGACCAGAATAGCAACAG 1690
Db       1409 AAGCAAAATTGCCTTGGCTGAAACAAGCACAAGAACTAGAAGAGACCAGAATAGCAACAG 1468

Qy       1691 AAGAACCAACGTTCATTCCAGAACCCTGGTCAGCCTGCAGTACCACGTGTGGGCCAGGTG 1750
Db       1469 AAGAACCAACGTTCATTCCAGAACCCTGGTCAGCCTGCAGTACCACGTGTGGGCCAGGTG 1528

Qy       1751 TGCAGGTCCGCGAGGTGAACTGCCGTGTGCTCCTCACATTCACGCAGACTGAGACTGAGC 1810
Db       1529 TGCAGGTCCGCGAGGTGAACTGCCGTGTGCTCCTCACATTCACGCAGACTGAGACTGAGC 1588

Qy       1811 TGCCCTAGGAAGACTTGGAAGGCCCCAAGCTGCCCACCGAACGGCCCTGCCTCCTGGAAG 1870
Db       1589 TGCCCTAGGAAGACTTGGAAGGCCCCAAGCTGCCCACCGAACGGCCCTGCCTCCTGGAAG 1648

Qy       1871 CATGTGATGAGAGCCCGGCCTCCCGAGAGCTAGACATCCCTCTCCCTGAGGACAGTGACA 1930
Db       1649 CATGTGATGAGAGCCCGGCCTCCCGAGAGCTAGACATCCCTCTCCCTGAGGACAGTGACA 1708

Qy       1931 CCACTTACGACTGGGAGTACCCTGGGTTCACCCCTTGCACAGCAACATGCCTTGGGAGTCC 1990
Db       1709 CCACTTACGACTGGGAGTACCCTGGGTTCACCCCTTGCACAGCAACATGCCTTGGGAGTCC 1768

Qy       1991 ATCAAGAAGCCATAGCAGTTTGCTTACATATCCAGACCCAGCAGACAGTCAATGACAGCT 2050
Db       1769 ATCAAGAAGCCATAGCAGTTTGCTTACATATCCAGACCCAGCAGACAGTCAATGACAGCT 1828

Qy       2051 TGTTGATATGGTCCACCGTCCTCCAGCCATGAGCCAGGCCTGTAACACAGAGCCCTGTC 2110
Db       1829 TGTTGATATGGTCCACCGTCCTCCAGCCATGAGCCAGGCCTGTAACACAGAGCCCTGTC 1888

Qy       2111 CCCCCAGGTGGCATGTGGGCTCTTGGGGGCCCTGCTCAGCTACCTGTGGAGTTGGAATTC 2170
Db       1889 CCCCCAGGTGGCATGTGGGCTCTTGGGGGCCCTGCTCAGCTACCTGTGGAGTTGGAATTC 1948

Qy       2171 AGACCCGAGATGTGTACTGCCTGCACCCAGGGGAGACCCCTGCCCCTGCCAGGAGTTCCC 2230
Db       1949 AGACCCGAGATGTGTACTGCCTGCACCCAGGGGAGACCCCTGCCCCTGCCAGGAGTTCCC 2008

Qy       2231 GAGATGAAAAGCCCCATGCTTTACAAGCATGCAATCAGTTTGACTGCCCCTGCCTGGCTTC 2290
Db       2009 GAGATGAAAAGCCCCATGCTTTACAAGCATGCAATCAGTTTGACTGCCCCTGCCTGGCTTC 2068

Qy       2291 ACATTGAAGAATGGCACCACCCCCCCAGGACTTGTGGCCGGGGGAACTCAGAACAGAAGAT 2350
Db       2069 ACATTGAAGAATGGCACCACTTCTCAGGACTTGTGGCCGGGGGAACTCAGAACAGAAGAT 2128

Qy       2351 TCACCTGTCCGCCAGCTGCTAACCGATGGCAGCTTTTTGAATCTCTCAGATGAATTGTGCC 241
Db       2129 TCACCTGTCCGCCAGCTGCTAACCGATGGCAGCTTTTTGAATCTCTCAGATGAATTGTGCC 2188

Qy       2411 AAGGACCCAAGGCATCGTCTCACAAGTCCTGTGCCAGGACAGACTGTCCTCCACATTTAG 2470
Db       2189 AAGGACCCAAGGCATCGTCTCACAAGTCCTGTGCCAGGACAGACTGTCCTCCACATTTAG 2248

Qy       2471 CTGTGGGAGACTGGTGGAAGTGTTCTGTCAGTTGCGGTCTTGGAATCCAGAGAAGAAAGC 2530
Db       2249 CTGTGGGAGACTGGTGGAAGTGTTCTGTCAGTTGCGGTCTTGGAATCCAGAGAAGAAAGC 2308

Qy       2531 AGGTGTGTCAAAGGCTGGCAGCTAAAGGTCGGCGCATGCCCCTCAGTGAGATGATGGTCA 2590
Db       2309 AGGTGTGTCAAAGGCTGGCAGCTAAAGGTCGGCGCATGCCCCTCAGTGAGATGATGGTCA 2368

Qy       2591 GGGATCTACCAGGGTTCCCTCTTGTAAGATCTTGCCCAGATGCCTGACTGCAGTAAAATCA 2650
Db       2369 GGGATCTACCAGGGTTCCCTCTTGTAAGATCTTGCCCAGATGCCTGACTGCAGTAAAATCA 2428

Qy       2651 AATCAGAGATGAAGACAAAACTTGGTGAGCAGGGTCCGCAGATCCTCAGGTGTCAGAGAG 2710
Db       2429 AATCAGAGATGAAGACAAAACTTGGTGAGCAGGGTCCGCAGATCCTCAGGTGTCAGAGAG 2488

Qy       2711 TCTACATTCAGACAAGGGAAGAGAAGCGTATTAACCTGACCATTGGTAGCAGAGCCTATT 2770
Db       2489 TCTACATTCAGACAAGGGAAGAGAAGCGTATTAACCTGACCATTGGTAGCAGAGCCTATT 2548

Qy       2771 TGCTGCCCAACACATCCGTGATTATTAAGTGCCCCGTGCGACGATTCCAGAAATCTCTGA 2830
Db       2549 TGCTGCCCAACACATCCGTGATTATTAAGTGCCCCGTGCGACGATTCCAGAAATCTCTGA 2608

Qy       2831 TCCAGTGGGAGAAGGATGGCCGTTGCCTGCAGAACTCCAAACGGCTTGGCATCACCAAGT 2890
Db       2609 TCCAGTGGGAGAAGGATGGCCGTTGCCTGCAGAACTCCAAACGGCTTGGCATCACCAAGT 2668

Qy       2891 CAGGCTCACTAAAAATCCACGGTCTTGCTGCCCCCGACATCGGCGTGTACCGGTGCATTG 2950
Db       2669 CAGGCTCACTAAAAATCCACGGTCTTGCTGCCCCCGACATCGGCGTGTACCGGTGCATTG 2728

Qy       2951 CAGGCTCTGCACAGGAAACAGTTGTGCTCAAGCTCATTGGTACTGACAACCGGCTCATCG 3010
Db       2729 CAGGCTCTGCACAGGAAACAGTTGTGCTCAAGCTCATTGGTACTGACAACCGGCTCATCG 2788

Qy       3011 CACGCCCAGCCCTCAGGGAGCCTATGAGGGAATATCCTGGGATGGACCACAGGCAAGCCA 3070
Db       2789 CACGCCCAGCCCTCAGGGAGCCTATGAGGGAATATCCTGGGATGGACCACAGGCAAGCCA 2848

Qy       3071 ATAGTTTGGGAGTCACATGGCACAAAATGAGGCAAATGTGGAATAACAAAAATGACCTTT 3130
Db       2849 ATAGTTTGGGAGTCACATGGCACAAAATGAGGCAAATGTGGAATAACAAAAATGACCTTT 2908

Qy       3131 ATCTGGATGATGACCACATTAGTAACCAGCCTTTCTTGAGAGCTCTGTTAGCCCACTGCA 3190
Db       2909 ATCTGGATGATGACCACATTAGTAACCAGCCTTTCTTGAGAGCTCTGTTAGCCCACTGCA 2968

Qy       3191 GCAATTCTGCAGGAAGCACCAACTCCTGGGAGTTGAAGAATAACCAGTTTCAAGCAGCAG 3250
Db       2969 GCAATTCTGCAGGAAGCACCAACTCCTGGGAGTTGAAGAATAACCAGTTTCAAGCAGCAG 3028

Qy       3251 TTAAACAAGGAGCATATAGCATGGATACAGCCCAGTTTGATGAGCTGATAAGAAACATGA 3310
Db       3029 TTAAACAAGGAGCATATAGCATGGATACAGCCCAGTTTGATGAGCTGATAAGAAACATGA 3088

Qy       3311 GTCAGCTCATGGAAACCGGAGAGGTCAGCGATGATCTTGCCGTCCCAGCTGATATATCAGC 3370
Db       3089 GTCAGCTCATGGAAACCGGAGAGGTCAGCGATGATCTTGCCGTCCCAGCTGATATATCAGC 3148

Qy       3371 TGGTGGCCGAATTAGCCAAGGCACAGCCAACACACATGCAGTTGGCGGCCATCCAGGAAG 3430
Db       3149 TGGTGGCCGAATTAGCCAAGGCACAGCCAACACACATGCAGTTGGCGGCCATCCAGGAAG 3208

Qy       3431 AGACACCTCCTGCTGCTCAGCTCAGAGGGGAAACAGGGAGGTGTGCCAAAGCTGGCATG 349
Db       3209 AGACACCTCCTGCTGCTCAGCTCAGAGGGGAAACAGGGAGGTGTGCCAAAGCTGGCATG 3268

Qy       3491 CAAAAAACTCAGGCAAGCTGACATTCAAGCCGAAAGGACCTGCTTCTCATGAGGTAAAGCC 3550
Db       3269 CAAAAAACTCAGGCAAGCTGACATTCAAGCCGAAAGGACCTGCTTCTCATGAGGTAAAGCC 3328

Qy       3551 AACCTCCCTCAATTTCATTTAATAAAACAATAAATTCCAGGATTGGAAATACAGTATACA 3610
Db       3329 AACCTCCCTCAATTTCATTTAATAAAACAATAAATTCCAGGATTGGAAATACAGTATACA 3388

Qy       3611 TTACAAAAAGGACAGAGGTCATCAATATACTGTGTGACCTTATTACCCCAGTGAGGCCA 3670
Db       3389 TTACAAAAAGGACAGAGGTCATCAATATACTGTGTGACCTTATTACCCCAGTGAGGCCA 3448

Qy       3671 CATATACATGGACCAAGGATGGAACCTTGTTACAGTTTCAGTAAAAATAATTTTGGATG 3730
```

```
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3449 CATATACATGGACCAAGGATGGAACCTTGTTACAGCCCTCAGTAAAAATAATTTTGGATG 3508

Qy       3731 GAACTGGGAAGATACAGATACAGAATCCTACAAGGAAAGAACAAGGCATATATGAATGTT 3790
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3509 GAACTGGGAAGATACAGATACAGAATCCTACAAGGAAAGAACAAGGCATATATGAATGTT 3568

Qy       3791 CTGTAGCTAATCATCTTGGTTCAGATGTGGAAAGTTCTTCTGTGCTGTATGCAGAGGCAC 3850
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3569 CTGTAGCTAATCATCTTGGTTCAGATGTGGAAAGTTCTTCTGTGCTGTATGCAGAGGCAC 3628

Qy       3851 CTGTCATCTTGTCTGTTGAAAGAAATATCACCAAACCAGAGCACAACCATCTGTCTGTTG 3910
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3629 CTGTCATCTTGTCTGTTGAAAGAAATATCACCAAACCAGAGCACAACCATCTGTCTGTTG 3688

Qy       3911 TGGTTGGAGGCATCGTGGAGGCAGCCCTTGGAGCAAACGTGACAATCCGATGTCCTGTAA 3970
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3689 TGGTTGGAGGCATCGTGGAGGCAGCCCTTGGAGCAAACGTGACAATCCGATGTCCTGTAA 3748

Qy       3971 AAGGTGTCCCTCAGCCTAATATAACTTGGTTGAAGAGAGGAGGATCTCTGAGTGGCAATG 4030
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3749 AAGGTGTCCCTCAGCCTAATATAACTTGGTTGAAGAGAGGAGGATCTCTGAGTGGCAATG 3808

Qy       4031 TTTCCTTGCTTTTCAATGGATCCCTGTTGTTGCAGAATGTTTCCCTTGAAAATGAAGGAA 4090
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3809 TTTCCTTGCTTTTCAATGGATCCCTGTTGTTGCAGAATGTTTCCCTTGAAAATGAAGGAA 3868

Qy       4091 CCTACGTCTGCATAGCCACCAATGCTCTTGGAAAGGCAGTGGCAACATCTGTACTCCACT 4150
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3869 CCTACGTCTGCATAGCCACCAATGCTCTTGGAAAGGCAGTGGCAACATCTGTACTCCACT 3928

Qy       4151 TGCTGGAACGAAGATGGCCAGAGAGTAGAATCGTATTTCTGCAAGCACATAAAAAGTACA 4210
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3929 TGCTGGAACGAAGATGGCCAGAGAGTAGAATCGTATTTCTGCAAGGACATAAAAAGTACA 3988

Qy       4211 TTCTCCAGGCAACCAACACTAGAACCAACAGCAATGACCCAACAGCACAACCCCCGCCTC 4270
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3989 TTCTCCAGGCAACCAACACTAGAACCAACAGCAATGACCCAACAGGACAACCCCCGCCTC 4048

Qy       4271 AAGAGCCTTTTTGCCAGCCTGGTAACTGGTCACATTGTTCTGCCACCTGTGGTCATTTGG 4330
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4049 AAGAGCCTTTTTGGGAGCCTGGTAACTGGTCACATTGTTCTGCCACCTGTGGTCATTTGG 4108

Qy       4331 GAGCCCGCATTCAGAGACCCCAGTGTGTGATGCCCAATGGGCAGGAACTGAGTGAGGCCC 4390
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4109 GAGCCCGCATTCAGAGACCCCAGTGTGTGATGGCCAATGGGCAGGAAGTGAGTGAGGCCC 4168

Qy       4391 TGTGTGATCACCTCCAGAAGCCACTGGCTGGGTTTGAGCCCTGTAACATCCGGGACTGCC 4450
              ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db       4169 TGTGTGATCACCTCCAGAAGCCACTGGCTGGGTTTGAGCCCTCTAACATCCGGGACTGCC 4228

Qy       4451 CAGCGAGGTGGTTCACAAGTGTGTGGTCACAGTGCTCTGTGTCTTGCGGTGAAGGATACC 4510
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4229 CAGCGAGGTGGTTCACAAGTGTGTGGTCACAGTGCTCTGTGTCTTGCGGTGAAGGATACC 4288

Qy       4511 ACAGTCGGCAGGTGACGTGCAAGCGGACAAAAGCCAATGGAACTGTGCAGGTGGTGTCTC 4570
              |||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db       4289 ACAGTCGGCAGGTGACGTGCAAGCGGACAAAAGCCAATGGAACTGTGCAGGTGGTGTCTC 4348

Qy       4571 CAAGAGCATGTGCCCCTAAAGACCGGCCTCTGGGAAGAAAACCATGTTTTGGTCATCCAT 4630
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4349 CAAGAGCATGTGCCCCTAAAGACCGGCCTCTGGGAAGAAAACCATGTTTTGGTCATCCAT 4408

Qy       4631 GTGTTCAGTGGGAACCAGGGAACCGGTGTCCTGGACCTTGCATGGGCCGTGCTGTGAGGA 4690
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4409 GTGTTCAGTGGGAACCAGGGAACCGGTGTCCTGGACCTTGCATGGGCCGTGCTGTGAGGA 4468

Qy       4691 TGCAGCAGCGTCACACAGCTTGTCAACACAACAGCTCTGACTCCAACTGTGATGACAGAA 4750
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4469 TGCAGCAGCGTCACACAGCTTGTCAACACAACAGCTCTCACTCCAACTGTGATGACAGAA 4528

Qy       4751 AGAGACCCACCTTAAGAAGGAACTGCACATCAGGGGCCTGTGATGTGTGTTGGCACACAG 4810
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4529 AGAGACCCACCTTAAGAAGGAACTGCACATCAGGGGCCTGTGATGTGTGTTGGCACACAG 4588

Qy       4811 GCCCTTGGAAGCCCTGTACAGCAGCCTGTGGCAGGGGTTTCCAGTCTCGGAAACTCGACT 4870
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4589 GCCCTTGGAAGCCCTGTACAGCAGCCTGTGGCAGGGGTTTCCAGTCTCGGAAACTCGACT 4648

Qy       4871 GTATCCACACAAGGAGTTGCAAACCTGTGGCCAAGAGACACTGTGTACAGAAAAAGAAAC 4930
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4649 GTATCCACACAAGGAGTTGCAAACCTGTGGCCAAGAGACACTGTGTACAGAAAAAGAAAC 4708

Qy       4931 CAATTTCCTGGCGGCACTGTCTTGGGCCCTCCTGTGATAGAGACTGCACAGACACAACTC 4990
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4709 CAATTTCCTGGCGGCACTGTCTTGGGCCCTCCTGTGATAGAGACTGCACAGACACAACTC 4768

Qy       4991 ACTACTGTATGTTTGTAAAACATCTTAATTTGTGTTCTCTAGACCGCTACAAACAAAGGT 5050
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4769 ACTACTGTATGTTTGTAAAACATCTTAATTTGTGTTCTCTAGACCGCTACAAACAAAGGT 4828

Qy       5051 GCTGCCAGTCATGTCAAGAGGGATAA 5076
              ||||||||||||||||||||||||||
Db       4829 GCTGCCAGTCATGTCAAGAGGGATAA 4854


RESULT 3
US-09-784-358-17
; Sequence 17, Application US/09784358
; Patent No. US20020099027A1
; GENERAL INFORMATION:
;  APPLICANT: Donoho, Gregory
;  APPLICANT:  Scoville, John
;  APPLICANT:  Turner, C. Alexander Jr.
;  APPLICANT:  Friedrich, Glenn
;  APPLICANT:  Zambrowicz, Brian
;  APPLICANT:  Sands, Arthur T.
;  TITLE OF INVENTION: NOVEL HUMAN THROMBOSPONDIN REPEAT PROTEINS AND
;  TITLE OF INVENTION:  POLYNUCLEOTIDES ENCODING THE SAME
;  FILE REFERENCE: LEX-0134-USA
;  CURRENT APPLICATION NUMBER: US/09/784,358
;  CURRENT FILING DATE:  2001-02-15
;  PRIOR APPLICATION NUMBER: US 60/183,282
;  PRIOR FILING DATE: 2000-02-17
;  NUMBER OF SEQ ID NOS: 17
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17
;   LENGTH: 8578
;   TYPE: DNA
;   ORGANISM: homo sapiens
US-09-784-358-17

  Query Match             74.0%;  Score 3755;  DB 10;  Length 8578;
  Best Local Similarity   79.7%;  Pred. No. 0;
  Matches 5076;  Conservative    0;  Mismatches    0;  Indels 1291;  Gaps    3;

Qy          1 ATGGCTTCCTGGACGAGCCCCTGGTGGGTGCTGATAGGGATGGTCTTCATGCACTCTCCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        179 ATGGCTTCCTGGACGAGCCCCTGGTGGGTGCTGATAGGGATGGTCTTCATGCACTCTCCC 238

Qy         61 CTCCCGCAGACCACAGCTGAGAAATCTCCTGGAGCCTATTTCCTTCCCGAGTTTGCACTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        239 CTCCCGCAGACCACAGCTGAGAAATCTCCTGGAGCCTATTTCCTTCCCGAGTTTGCACTT 298

Qy        121 TCTCCTCAGGGAAGTTTTCTGGAAGACACAACAGGGGAGCAGTTCCTCACTTATCGCTAT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        299 TCTCCTCAGGGAAGTTTTCTGGAAGACACAACAGGGGAGCAGTTCCTCACTTATCGCTAT 358

Qy        181 GATGACCAGACCTCAAGAAACACTCGTTCAGATGAAGACAAAGATGGCAACTGGGATGCT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        359 GATGACCAGACCTCAAGAAACACTCGTTCAGATGAAGACAAAGATGGCAACTGGGATGCT 418

Qy        241 TGGGGCGACTGGAGTGACTGCTCCCGGACCTGTGGGGGAGGAGCATCATATTCTCTGCGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        419 TGGGGCGACTGGAGTGACTGCTCCCGGACCTGTGGGGGAGGAGCATCATATTCTCTGCGG 478
```

```
Qy        301 AGATGTTTGACTGGAAGGAATTGTGAAGGGCAGAACATTCGGTACAAGACATGCAGCAAT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        479 AGATGTTTGACTGGAAGGAATTGTGAAGGGCAGAACATTCGGTACAAGACATGCAGCAAT 538

Qy        361 CATGACTGCCCTGCAGATCCAGAAGATTTCAGAGCCAGCAGGGCTCAGCCTACAATGAT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        539 CATGACTGCCCTGCAGATCCAGAAGATTTCAGAGCCAGCAGGGCTCAGCCTACAATGAT 598

Qy        421 GTCCAGTATCAGGGGCATTACTATGAATGGCTTCCACGATATAATGATCCTGCTGCCCG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        599 GTCCAGTATCAGGGGCATTACTATGAATGGCTTCCACGATATAATGATCCTGCTGCCCG 658

Qy        481 TGTGCACTCAAGGGTCATGCACAAGGACAAAACTTGGTGGTGGAGCTGGCACCTAAGGTA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        659 TGTGCACTCAAGGGTCATGCACAAGGACAAAACTTGGTGGTGGAGCTGGCACCTAAGGTA 718

Qy        541 CTGGATGGAACTGGTTGCAACAGGGACTGCTTGGACATGGTGATCAGTGGCATCTGTCAG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        719 CTGGATGGAACTGGTTGCAACAGGGACTGCTTGGACATGGTGATCAGTGGCATCTGTCAG 778

Qy        601 GCAGTGGGCTGGGATGGGCAACTGGGAAGCAATGCCAAGGAGGACAACTGGTGGAGTCTGT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        779 GCAGTGGGCTGGGATGGGCAACTGGGAAGCAATGCCAAGGAGGACAACTGGTGGAGTCTGT 838

Qy        661 GCCGGGGATGGCTGCACCTGCAGGCTTGTACGGGACAATCAAAGTCACAGTCTTCTCCTT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        839 GCCGGGGATGGCTGCACCTGCAGGCTTGTACGGGACAATCAAAGTCACAGTCTTCTCCTT 898

Qy        721 GAAAAAAGAGAAGAAAATGTAATTGCTGTTCCTTTGGGAAGTCGAAGTCTGAGAATTACA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        899 GAAAAAAGAGAAGAAAATGTAATTGCTGTTCCTTTGGGAAGTCGAAGTCTGAGAATTACA 958

Qy        781 GTGAAAGGACCTGCCCACCTCTTTATTGAATCAAAAACACTTCAAGGAAGCAAAGGAGAA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        959 GTGAAAGGACCTGCCCACCTCTTTATTGAATCAAAAACACTTCAAGGAAGCAAAGGAGAA 1018

Qy        841 CACAGCTTTAACAGCCCCGGCGTCTTTGTCGTAGAAAACACAACAGTGGAATTTCAGAGG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1019 CACAGCTTTAACAGCCCCGGCGTCTTTGTCGTAGAAAACACAACAGTGGAATTTCAGAGG 1078

Qy        901 GGCTCCGAGAGGCAAACTTTTAAGATTCCAGGACCTCTGATGGCTGATTTCATCTTCAAT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1079 GGCTCCGAGAGGCAAACTTTTAAGATTCCAGGACCTCTGATGGCTGATTTCATCTTCAAT 1138

Qy        961 ACCAGGTACACTGCAGCCAAAGACAGCGTGGTTCAGTTCTTCTTTTACCAGCCCATCAGT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1139 ACCAGGTACACTGCAGCCAAAGACAGCGTGGTTCAGTTCTTCTTTTACCAGCCCATCAGT 1198

Qy       1021 CATCAGTGGAGACAAACTGACTTCTTTTCCCTGCACTGTGACGTGTGGAGGAGGTTATCAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1199 CATCAGTGGAGACAAACTGACTTCTTTTCCCTGCACTGTGACGTGTGGAGGAGGTTATCAG 1258

Qy       1081 CTCAATTCTGCTGAATGTGTGGATATCCGCTTGAAGAGGGTAGTTCCTGACCATTATTGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1259 CTCAATTCTGCTGAATGTGTGGATATCCGCTTGAAGAGGGTAGTTCCTGACCATTATTGT 1318

Qy       1141 CACTACTACCCTGAAAATGTAAAACCAAAACCAAAACTGAAGGAATGCAGCATGGATCCC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1319 CACTACTACCCTGAAAATGTAAAACCAAAACCAAAACTGAAGGAATGCAGCATGGATCCC 1378

Qy       1201 TGCCCATCAAGTGATGGATTTAAAGAGATAATGCCCTATGACCACTTCCAACCTCTTCCT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1379 TGCCCATCAAGTGATGGATTTAAAGAGATAATGCCCTATGACCACTTCCAACCTCTTCCT 1438

Qy       1261 C--GCTGGGAACATAAGCCTGGACTGCATGTTCCGTGTCCTGTGGAGGAGGCATTCAGA 1318
              |  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1439 CCAGCTGGGAACATAAGCCTGGACTGCATGTTCCGTGTCCTGTGGAGGAGGCATTCAGA 1498

Qy       1319 GACGGAGCTTTGTGTGTGTAGAGGAATCCATCCATGGAGAGATATTGCAGTGGAAGAAT 1378
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1499 GACGGAGCTTTGTGTGTGTAGAGGAATCCATCCATGGAGAGATATTGCAGTGGAAGAAT 1558

Qy       1379 GGAAGTGCATGTACGCACC-AAACCAAGGTTATGAAACTTGTAATCTGTTTGATTGCC 1438
              ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db       1559 GGAAGTGCATGTACGCACC-AAACCAAGGTTATGAAACTTGTAATCTGTTTGATTGCC 1618

Qy       1439 CCAAGTGGATTGCCATGGAGTGGTTCAGTGCACAGTGACTTGTGCCGAGGGTTACCGT 1498
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1619 CCAAGTGGATTGCCATGGAGTGGTTCAGTGCACAGTGACTTGTGCCGAGGGTTACCGT 1678

Qy       1499 ACCGGGTTGTTCTGTGTATTAACCACCGCGGAGAGCATGTTGGGGGCTGCAATCCACAAC 1558
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1679 ACCGGGTTGTTCTGTGTATTAACCACCGCGGAGAGCATGTTGGGGGCTGCAATCCACAAC 1738

Qy       1559 TGAAGTTACACATCAAAGAAGAATGTGCATTCCCATCCTGGGTATAAACCAAAAGAAA 1618
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1739 TGAAGTTACACATCAAAGAAGAATGTGCATTCCCATCCTGGGTATAAACCAAAAGAAA 1798

Qy       1619 AAAGTCCAGTGGAAGCAAAATTGCCTTGGCTGAAACAAGCACAAGAATTAGAAGAGACCA 1678
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1799 AAAGTCCAGTGGAAGCAAAATTGCCTTGGCTGAAACAAGCACAAGAATTAGAAGAGACCA 1858

Qy       1679 GAATAGCAACAGAAGAACCAACTGTCATTCCAGAACCCTGCCTACCCTGCATTACCAGT 1738
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1859 GAATAGCAACAGAAGAACCAACTGTCATTCCAGAACCCTGCCTACCCTGCATTACCAGT 1918

Qy       1739 GTGGGCCAGGTGTGCAGGTCCCGGAGGTGAAGTGCCGTGTGCTCCTCACATTCACGCAGA 1798
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1919 GTGGGCCAGGTGTGCAGGTCCCGGAGGTGAAGTGCCGTGTGCTCCTCACATTCACGCAGA 1978

Qy       1799 CTGAGACTGAGCTGGCCTGAGGAAGAGGTGAAGGCCCCAAGCTGCCCACCGAAGGACCCT 1858
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1979 CTGAGACTGAGCTGGCCTGAGGAAGAGGTGAAGGCCCCAAGCTGCCCACCGAAGGACCCT 2038

Qy       1859 GCCTCCTGGAAGCATGTGATGAGAGCCCGGCCTCCCGAGAGCTAGACATCCCTCTCCTG 1918
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2039 GCCTCCTGGAAGCATGTGATGAGAGCCCGGCCTCCCGAGAGCTAGACATCCCTCTCCTG 2098

Qy       1919 AGGACAGTGAGACGACTTACGACTGGGAGTACGCTGGGTTCACCCCTTGCACAGCAACAT 1978
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2099 AGGACAGTGAGACGACTTACGACTGGGAGTACGCTGGGTTCACCCCTTGCACAGCAACAT 2158

Qy       1979 GCTTGGGAGGCCATCAAGAAGCCATAGCAGTGTGCTTACATATCCAGACCCAGCAGACAG 2038
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2159 GCTTGGGAGGCCATCAAGAAGCCATAGCAGTGTGCTTACATATCCAGACCCAGCAGACAG 2218

Qy       2039 TCAATGACAGCTTGTGTGATATGGTCCACCGTCCTCCAGCCATGAGCCAGGCCTGTAACA 2098
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2219 TCAATGACAGCTTGTGTGATATGGTCCACCGTCCTCCAGCCATGAGCCAGGCCTGTAACA 2278

Qy       2099 CAGAGCCCTGTCCCCCAGT                                          2118
              ||||||||||||||||||||
Db       2279 CAGAGCCCTGTCCCCCAGGAGAGAGCCAGCAGCTTGTAGAAGCATCCGGGTTACATAA 2338

Qy       2119                                                              2118

Db       2339 TGGTCCTGCCTAGTCTGAGGAGAGCCTTCTTCTCTAACAGGATTCAACACTGGCAGGGAAG 2398

Qy       2119                                                              2118

Db       2399 AAAGGAGGAAAGGAAGAGGCAATAGTGATGTGTTTCTGTACCAGCTTGTTACCTATTTCT 2458

Qy       2119                                                              2118

Db       2459 TGATATAAAAAACAATTCTTCTATTGAGTTCATTGTCTGTGAATAAGAAATTGTTGCCCAT 2518

Qy       2119                                           GGCCATGTGGGCTCT 2133
                                                        |||||||||||||||
Db       2519 TTCTTAAATAAAAACAGCCCCATCTCAAAAAAAAAAAAAAAAAAAGGCATGTGGGCTCT 2578

Qy       2134 TGGGGGCCCTGGCTCAGCTACCTGTGGAGTTGGAATTCAGACCCCAGATGTGTACTGGCCTG 2193
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2579 TGGGGGCCCTGGCTCAGCTACCTGTGGAGTTGGAATTCAGACCCCAGATGTGTACTGGCCTG 2638

Qy       2194 CACCCAGGGGGACCCCTGCCCCCTCCTGGAGGAGCCCCCAGAGGAAAAGCCCCATGCTTA 2253
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2639 CACCCAGGGGAGACCCCTGCCCCTCCTGAGGAGTGCCGAGATGAAAAGCCCCATGCTTTA 2698

Qy   2254 CAAGCATGCAATCAGTTTGACTGCCCTCCTGGCTGGCACATTGAAGAATGGCAGCAGTGT 2313
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2699 CAAGCATGCAATCAGTTTGACTGCCCTCCTGGCTGGCACATTGAAGAATGGCAGCAGTGT 2758

Qy   2314 TCCAGGACTTGTGGCGGGGGAACTCAGAACAGAAGAGTCACCTGTCGGCAGCTGCTAACG 2373
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2759 TCCAGGACTTGTGGCGGGGGAACTCAGAACAGAAGAGTCACCTGTCGGCAGCTGCTAACG 2818

Qy   2374 GATGGCAGCTTTTTGAATCTCTCAGATGAATTGTGCCAAGGACCCAAGGCATCGTCTCAC 2433
          |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db   2819 GATGGCAGCTTTTTGAATCTCTCAGATGAATTGTGCCAAGGACCCAAGGCATCGTCTCAC 2878

Qy   2434 AAGTCCTGTGCCAGGACAGACTGTCCTCCACATTTAGCTGTGGGAGACTGGTCGAAG--- 2490
          ||||||||||||||||||||||||||||||||||||.||||||||||||||||||||
Db   2879 AAGTCCTGTGCCAGGACAGACTGTCCTCCACATTTAGCTGTGGGAGACTGCTCGAAGGAG 2938

Qy   2491 ------------------------------------------------------------ 2490

Db   2939 CATTCAATGCAAGATTACAATCCAGCAGGATCTACACAATTCTAAAGAAAAGCAAGCATG 2998

Qy   2491 ------------------------------------------------------------ 2490

Db   2999 ACTCAAGGATTTCCTCTTCATCCTGTGTTCTTCATGTAGAGAGACAGCAGAGAGGCACTC 3058

Qy   2491 ------------------------------------------------------------ 2490

Db   3059 AGAGAATACTGTCTGATAAGCCCTTGAAAAAGCTGTAGGGCCAAGATGAGATACAGAGAT 3118

Qy   2491 ------------------------------------------------------------ 2490

Db   3119 GACTCAAAACAGAGAATCCAGGAATGCATAGATCCTGGTAAAAAAGGTGGAGATCACTA 3178

Qy   2491 ------------------------------------------------------------ 2490

Db   3179 ATAAATTCATTTGTGTAGGATTAAGACTAATCAACTAACAATTATATTATAGAACATAAC 3238

Qy   2491 ------------------------------------------------------------ 2490

Db   3239 ATAAATATCAGAAATTTTGACATTATCTAAATAATAAAATGAAAACTAATTGAGATTTGG 3298

Qy   2491 ------------------------------------------------------------ 2490

Db   3299 AGAGATGAGGTAGATGATATAGTTTGGCTGTGTCCCACCCAAATCTCATCTTGAATTGT 3358

Qy   2491 ------------------------------------------------------------ 2490

Db   3359 AGTTCCCATAATTCCCATGTGTTGTGGGAGGGACTCAGTTGGAGATAATTGAATCATGGG 3418

Qy   2491 ------------------------------------------------------------ 2490

Db   3419 GGCAGTTTCCCTCATACTGTTCTCGTGGTGGTAAATGAGTCTCACGAGATCTGATGGTTT 3478

Qy   2491 ------------------------------------------------------------ 2490

Db   3479 TATAAGGGGTTTCCCTTTTCGCTTGGCTCTCATTCTCTCTTGCCTGCTGCCATGTAAGAC 3538

Qy   2491 ------------------------------------------------------------ 2490

Db   3539 GTCCCTTTGCCCTTCCTTTGTCTTCTGCCATGATTGTGAGGCTTCCCTAGCCACGTGGAA 3598

Qy   2491 ------------------------------------------------------------ 2490

Db   3599 CTGAGTCTATTAAAGCTCTTTCCTTTATAACTTACCCAGTCTTGGGTATGTCTTTATTAA 3658

Qy   2491 ------------------------------------------------------------ 2490

Db   3659 CAACATAAGATTGGACTAATACAGTAGAGGAAATGTAAGTGTGCTTATTTCCTCATCCTT 3718

Qy   2491 ------------------------------------------------------------ 2490

Db   3719 CTTAGTAGCAAGTCAATAAATACTCTCCTAAGTCAAATTGTCATTAAAAATAACTATCCA 3778

Qy   2491 ------------------------------------------------------------ 2490

Db   3779 AATCTCTTGTTGGTTTATTTAATCTTCTTTATTAACTTTAGAGTGTTCTTTCGGGAATTA 3838

Qy   2491 ------------------------------------------------------------ 2490

Db   3839 ATCATGGTTTAAAAAATATCAAACATTCAACAACTCTAATTTTACTTTAATGTCTTTTTT 3898

Qy   2491 ------------------------------------------------------------ 2490

Db   3899 TCTAATATATCTAATAAAATTGCATTAAATTTTTAAGTTGAAAAAAAAAAAAAAAAAAAA 3958

Qy   2491 -TGTTCTGTCAGTTGTGGTGTTGGAATCCAGAGAAGAAAGCAGGTGTGTCAAAGGCTGGC 2549
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3959 ATGTTCTGTCAGTTGTGGTGTTGGAATCCAGAGAAGAAAGCAGGTGTGTCAAAGGCTGGC 4018

Qy   2550 AGCCAAAGGTCGGCGCATCCCCCTCAGTGAGATGATGTGCAGGGATCTACCAGGCTTCCC 2609
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4019 AGCCAAAGGTCGGCGCATCCCCCTCAGTGAGATGATGTGCAGGGATCTACCAGGCTTCCC 4078

Qy   2610 TCTTGTAAGATCTTGCCAGATGCCTGAGTGCAGTAAAATCAAATCAGAGATGAAGACAAA 2669
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4079 TCTTGTAAGATCTTGCCAGATGCCTGAGTGCAGTAAAATCAAATCAGAGATGAAGACAAA 4138

Qy   2670 ACTTGGTGAGCAGGGTCCGCAGATCCTCAGTGTCCAGAGAGTCTACATTCAGACAAGGGA 2729
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4139 ACTTGGTGAGCAGGGTCCGCAGATCCTCAGTGTCCAGAGAGTCTACATTCAGACAAGGGA 4198

Qy   2730 AGAGAAGCGTATTAACCTGACCATTGGTAGCAGAGCCTATTTGCTGCCCAACACATCCGT 2789
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4199 AGAGAAGCGTATTAACCTGACCATTGGTAGCAGAGCCTATTTGCTGCCCAACACATCCGT 4258

Qy   2790 GATTATTAAGTGCCCCGTGCGACGATTCCAGAAATCTCTGATCCAGTGGGAGAAGGATGG 2849
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4259 GATTATTAAGTGCCCCGTGCGACGATTCCAGAAATCTCTGATCCAGTGGGAGAAGGATGG 4318

Qy   2850 CCGTTGCCTGCAGAACTCCAAACGGCTTGGCATCACCAAGTCAGGCTCACTAAAAATCCA 2909
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4319 CCGTTGCCTGCAGAACTCCAAACGGCTTGGCATCACCAAGTCAGGCTCACTAAAAATCCA 4378

Qy   2910 CGGTCTTGCTGCCCCCGACATCGGCGTGTACCGGTGCATTGCAGGCTCTGCACAGGAAAC 2969
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4379 CGGTCTTGCTGCCCCCGACATCGGCGTGTACCGGTGCATTGCAGGCTCTGCACAGGAAAC 4438

Qy   2970 AGTTGTGCTCAAGCTCATTGGTACTGACAACCGGCTCATCGCACGCCCAGCCCTCAGGGA 3029
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4439 AGTTGTGCTCAAGCTCATTGGTACTGACAACCGGCTCATCGCACGCCCAGCCCTCAGGGA 4498

Qy   3030 GCCTATGAGGGAATATCCTGGGATGGACCACAGCGAAGCCAATAGTTTGGGAGTCACATG 3089
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4499 GCCTATGAGGGAATATCCTGGGATGGACCACAGCGAAGCCAATAGTTTGGGAGTCACATG 4558

Qy   3090 GCACAAAATGAGGCAAATGTGGAATAACAAAAATGACCTTTATCTGGATGATGACCACAT 3149
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4559 GCACAAAATGAGGCAAATGTGGAATAACAAAAATGACCTTTATCTGGATGATGACCACAT 4618

Qy   3150 TAGTAACCAGCCTTTCTTGAGAGCTCTGTTAGGCCACTGCAGCAATTCTGCAGGAAGCAC 3209
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4619 TAGTAACCAGCCTTTCTTGAGAGCTCTGTTAGGCCACTGCAGCAATTCTGCAGGAAGCAC 4678

Qy   3210 CAACTCCTGGGAGTTGAAGAATAAGCAGTTTGAAGCAGCAGTTAAACAAGGAGCATATAG 3269
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4679 CAACTCCTGGGAGTTGAAGAATAAGCAGTTTGAAGCAGCAGTTAAACAAGGAGCATATAG 4738

Qy   3270 CATGGATACAGCCCAGTTTGATGAGCTGATAAGAAACATGAGTCAGCTCATGGAAACCGG 3329
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4739 CATGGATACAGCCCAGTTTGATGAGCTGATAAGAAACATGAGTCAGCTCATGGAAACCGG 4798

Qy   3330 AGAGGTCAGCGATGATCTTGCGTCCCAGCTGATATATCAGCTGGTGGCCGAATTAGCCAA 3389
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4799 AGAGGTCAGCGATGATCTTGCGTCCCAGCTGATATATCAGCTGGTGGCCGAATTAGCCAA 4858
```

```
Qy       3390 GGCACAGCCAACACACATGCAGTGGGGGGCATCCAGGAAGAGACACCTCCCTGCTGCTCA 3449
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4859 GGCACAGCCAACACACATGCAGTGGGGGGCATCCAGGAAGAGACACCTCCCTGCTGCTCA 4918

Qy       3450 GCTCAGAGGGGAAACAGGGAGTGTGTCCCAAAGCTCGGCATGCAAAAACTCAGGCAAGCT 3509
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4919 GCTCAGAGGGGAAACAGGGAGTGTGTCCCAAAGCTCGGCATGCAAAAACTCAGGCAAGCT 4978

Qy       3510 GACATTCAAGCCGAAAGGACTTGTTCTCATGAGGCAAAGCCAACCTCCCTCAATTTCATT 3569
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4979 GACATTCAAGCCGAAAGGACTTGTTCTCATGAGGCAAAGCCAACCTCCCTCAATTTCATT 5038

Qy       3570 TAATAAAACAATAAATTCTACTGATTGGAAATACAGTATACATTACAAAAAGACAGAGTT 3629
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5039 TAATAAAACAATAAATTCTACTGATTGGAAATACAGTATACATTACAAAAAGACAGAGTT 5098

Qy       3630 CATCAATATACTGGGTGGACCTTATTACCCCAGTGAGGCCACATATACATGGACCAAGCA 3689
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5099 CATCAATATACTGGGTGGACCTTATTACCCCAGTGAGGCCACATATACATGGACCAAGCA 5158

Qy       3690 TGGAACCTTGTTACAGCCTCAGTAAAAAATAATTTTGGATGGAACTGGGAAGATACAGAT 3749
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5159 TGGAACCTTGTTACAGCCTCAGTAAAAAATAATTTTGGATGGAACTGGGAAGATACAGAT 5218

Qy       3750 ACAGAATCCTACAAGGAAAGAACAAGGCATATATGAATGTTCTGTAGCTAATCATCTTGG 3809
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5219 ACAGAATCCTACAAGGAAAGAACAAGGCATATATGAATGTTCTGTAGCTAATCATCTTGG 5278

Qy       3810 TTCAGATGTGGAAAGTTCTTCTGTGCTGTATGCAGAGGCACCTGTCATCTTGTCTGTTGA 3869
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5279 TTCAGATGTGGAAAGTTCTTCTGTGCTGTATGCAGAGGCACCTGTCATCTTGTCTGTTGA 5338

Qy       3870 AAGAAATATCACCAAACCAGAGCACAACCATCTGTCTGTTGTGGTTGGAGGCATCGTGGA 3929
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5339 AAGAAATATCACCAAACCAGAGCACAACCATCTGTCTGTTGTGGTTGGAGGCATCGTGGA 5398

Qy       3930 GGCAGCCCTTGGAGCAAACGTGACAATCCGATGTCCTGTAAAAGGTGTCCCTCAGCCTAA 3989
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5399 GGCAGCCCTTGGAGCAAACGTGACAATCCGATGTCCTGTAAAAGGTGTCCCTCAGCCTAA 5458

Qy       3990 TATAACTTGGTTGAAGAGACGAGGATCTCTGAGTGGCAATGTTTCCTTGCTTTTCAATCC 4049
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5459 TATAACTTGGTTGAAGAGACGAGGATCTCTGAGTGGCAATGTTTCCTTGCTTTTCAATCC 5518

Qy       4050 ATCCCTGGTTGTTGCAGAATGTTTCCTTGAAAATGAAGGAACCTACGTCTGCATAGCCAC 4109
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5519 ATCCCTGGTTGTTGCAGAATGTTTCCTTGAAAATGAAGGAACCTACGTCTGCATAGCCAC 5578

Qy       4110 CAATGCTGCTTGGAAAGGCAGTGGCAACATCTGTACTCCACTTGCTGGAACGAAGATGCC 4169
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5579 CAATGCTGCTTGGAAAGGCAGTGGCAACATCTGTACTCCACTTGCTGGAACGAAGATGCC 5638

Qy       4170 AGAGAGTAGAATGGTATTTCTGCAAGGACATAAAAAGTACATTCTCCAGGCAACCAACAC 4229
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5639 AGAGAGTAGAATGGTATTTCTGCAAGGACATAAAAAGTACATTCTCCAGGCAACCAACAC 5698

Qy       4230 TAGAACCAACAGCAATGACCCAACAGGAGAACCCCGCCTCAAGAGCCTTTTTCGGAACCC 4289
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5699 TAGAACCAACAGCAATGACCCAACAGGAGAACCCCGCCTCAAGAGCCTTTTTCGGAACCC 5758

Qy       4290 TGGTAACTGGTCACATTGTTTCTGCCACCTGTGGTCATTTGGGAGCCGCATTCAGAGACC 4349
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5759 TGGTAACTGGTCACATTGTTTCTGCCACCTGTGGTCATTTGGGAGCCGCATTCAGAGACC 5818

Qy       4350 CCAGTGTGTGATGGCCAATGGGCAGGAAGTGAGTGAGGCCTGTGTGATCACCTTCCAGAA 4409
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5819 CCAGTGTGTGATGGCCAATGGGCAGGAAGTGAGTGAGGCCTGTGTGATCACCTTCCAGAA 5878

Qy       4410 GCCACTGGCTGGGTTTGAGCCCTGTAACATCCGGACTGCCCAGCGAGGTTGGTTCACAAG 4469
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5879 GCCACTGGCTGGGTTTGAGCCCTGTAACATCCGGACTGCCCAGCGAGGTTGGTTCACAAG 5938

Qy       4470 TGTGTGGTCACAGTGCTCTGTCTCCCGGGTGAAGATACCACAGCCCCAGGTGACCGGCCC 4529
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5939 TGTGTGGTCACAGTGCTCTGTCTCCCGGGTGAAGATACCACAGCCCCAGGTGACCGGCCC 5998

Qy       4530 CAAGCGGACAAAAGCCAATGGAACTGTGCAGGTGGTGTCTCCAAGAGCATGTGCCCCTAA 4589
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5999 CAAGCGGACAAAAGCCAATGGAACTGTGCAGGTGGTGTCTCCAAGAGCATGTGCCCCTAA 6058

Qy       4590 AGACCGGCCTCTGGGAAGAAAACCATGTTTTTGGTCATCCATGGTTCAGTGGGAACCAGG 4649
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6059 AGACCGGCCTCTGGGAAGAAAACCATGTTTTTGGTCATCCATGGTTCAGTGGGAACCAGG 6118

Qy       4650 GAACCGGTGTCCTGGACGTTGCATGGGCCGTGCTGCAGGATGCAGCAGCGTCACACAGC 4709
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6119 GAACCGGTGTCCTGGACGTTGCATGGGCCGTGCTGCAGGATGCAGCAGCGTCACACAGC 6178

Qy       4710 TTGTCAACACAACAGCTCTGACTCCAACTGTGATGACAGAAAGAGACCCACCTTAAGAAG 4769
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6179 TTGTCAACACAACAGCTCTGACTCCAACTGTGATGACAGAAAGAGACCCACCTTAAGAAG 6238

Qy       4770 GAACTGACACATCAGGGGCCTGGGATGTGTGTTGGCACACAGCCCTTGGAAGCCCTGTAC 4829
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6239 GAACTGACACATCAGGGGCCTGGGATGTGTGTTGGCACACAGCCCTTGGAAGCCCTGTAC 6298

Qy       4830 AGCAGCCTGTGGGCAGGGGTTCCAGTCTCGGAAAGTCGACTGTATCCACACAAGGAGTTG 4889
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6299 AGCAGCCTGTGGGCAGGGGTTCCAGTCTCGGAAAGTCGACTGTATCCACACAAGGAGTTG 6358

Qy       4890 CAAACCTGTGGCCAAGAGACACTGTGTACAGAAAAAGAAACCAATTTCCTGGCGGCACTG 4949
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6359 CAAACCTGTGGCCAAGAGACACTGTGTACAGAAAAAGAAACCAATTTCCTGGCGGCACTG 6418

Qy       4950 TCTTGGGCCCTCCTGTGATAGAGACTGCACAGACACAACTCACTACTGTATGTTTGTAAA 5009
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6419 TCTTGGGCCCTCCTGTGATAGAGACTGCACAGACACAACTCACTACTGTATGTTTGTAAA 6478

Qy       5010 ACATCTTAATTTGTGTTCTCTAGACCGCTACAAACAAAGGTGCTGCCAGTCATGTCAAGA 5069
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6479 ACATCTTAATTTGTGTTCTCTAGACCGCTACAAACAAAGGTGCTGCCAGTCATGTCAAGA 6538

Qy       5070 GGGATAA 5076
              |||||||
Db       6539 GGGATAA 6545


RESULT 4
US-09-784-358-11
; Sequence 11, Application US/09784358
; Patent No. US20020090027A1
; GENERAL INFORMATION:
;  APPLICANT: Donoho, Gregory
;  APPLICANT:  Scoville, John
;  APPLICANT:  Turner, C. Alexander Jr.
;  APPLICANT:  Friedrich, Glenn
;  APPLICANT:  Zambrowicz, Brian
;  APPLICANT:  Sands, Arthur T.
;  TITLE OF INVENTION: NOVEL HUMAN THROMBOSPONDIN REPEAT PROTEINS AND
;  TITLE OF INVENTION:  POLYNUCLEOTIDES ENCODING THE SAME
;  FILE REFERENCE: LEX-0134-USA
;  CURRENT APPLICATION NUMBER: US/09/784,358
;  CURRENT FILING DATE:  2001-02-15
;  PRIOR APPLICATION NUMBER: US 60/183,28
;  PRIOR FILING DATE: 2000-02-17
;  NUMBER OF SEQ ID NOS: 17
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11
;   LENGTH: 2538
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-784-358-11

  Query Match             49.1%;  Score 2490;  DB 16;  Length 2538;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 2490;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          1 ATGGCTTCCTGGACGAGCCCCTGGTGGGTGCTGATAGGGATGGTCTTCATGCACTCTCCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGCTTCCTGGACGAGCCCCTGGTGGGTGCTGATAGGGATGGTCTTCATGCACTCTCCC 60

Qy         61 CTCCCGCAGACCACAGCTGAGAAATCTCCTGGAGCCTATTTCCTTCCCGAGTTTGCACTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTCCCGCAGACCACAGCTGAGAAATCTCCTGGAGCCTATTTCCTTCCCGAGTTTGCACTT 120

Qy        121 TCTCCTCAGGGAAGTTTTCTGGAAGACACAACAGGGGAGCAGTTCCTCACTTATCGCTAT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCTCCTCAGGGAAGTTTTCTGGAAGACACAACAGGGGAGCAGTTCCTCACTTATCGCTAT 180

Qy        181 GATGACCAGACCTCAAGAAACACTCGTTCAGATGAAGACAAAGATGGCAACTGGGATGCT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GATGACCAGACCTCAAGAAACACTCGTTCAGATGAAGACAAAGATGGCAACTGGGATGCT 240

Qy        241 TGGGGCGACTGGAGTGACTGCTCCCGGACCTGTGGGGGAGGAGCATCATATTCTCTGCGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGGGGCGACTGGAGTGACTGCTCCCGGACCTGTGGGGGAGGAGCATCATATTCTCTGCGG 300

Qy        301 AGATGTTTGACTGGAAGGAATTGTGAAGGGCAGAACATTCGGTACAAGACATGCAGCAAT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGATGTTTGACTGGAAGGAATTGTGAAGGGCAGAACATTCGGTACAAGACATGCAGCAAT 360

Qy        361 CATGACTGCCCTCCAGATGCAGAAGATTTCAGAGCCCAGCAGTGCTCAGCCTACAATGAT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CATGACTGCCCTCCAGATGCAGAAGATTTCAGAGCCCAGCAGTGCTCAGCCTACAATGAT 420

Qy        421 GTCCAGTATCAGGGGCATTACTATGAATGGCTTCCACGATATAATGATCCTGCTGCCCCG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GTCCAGTATCAGGGGCATTACTATGAATGGCTTCCACGATATAATGATCCTGCTGCCCCG 480

Qy        481 TGTGCACTCAAGTGTCATGCACAAGGACAAAACTTGGTGGTGGAGCTGGCACCTAAGGTA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TGTGCACTCAAGTGTCATGCACAAGGACAAAACTTGGTGGTGGAGCTGGCACCTAAGGTA 540

Qy        541 CTGGATGGAACTCGTTGCAACACGGACTCCTTGGACATGTGTATCAGTGGCATCTGTCAG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTGGATGGAACTCGTTGCAACACGGACTCCTTGGACATGTGTATCAGTGGCATCTGTCAG 600

Qy        601 GCAGTGGGCTGCGATCGGCAACTGGGAAGCAATGCCAAGGAGGACAACTGTGGAGTCTGT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GCAGTGGGCTGCGATCGGCAACTGGGAAGCAATGCCAAGGAGGACAACTGTGGAGTCTGT 660

Qy        661 GCCGGCGATGGCTCCACCTGCAGGCTTGTACGGGGACAATCAAAGTCACACGTTTCTCCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GCCGGCGATGGCTCCACCTGCAGGCTTGTACGGGGACAATCAAAGTCACACGTTTCTCCT 720

Qy        721 GAAAAAAGAGAAGAAAATGTAATTGCTGTTCCTTTGGGAAGTCGAAGTGTGAGAATTACA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GAAAAAAGAGAAGAAAATGTAATTGCTGTTCCTTTGGGAAGTCGAAGTGTGAGAATTACA 780

Qy        781 GTGAAAGGACCTGCCCACCTCTTTATTGAATCAAAAACACTTCAAGGAAGCAAAGGAGAA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GTGAAAGGACCTGCCCACCTCTTTATTGAATCAAAAACACTTCAAGGAAGCAAAGGAGAA 840

Qy        841 CACAGCTTTAACAGCCCCGGCGTCTTTGTCGTAGAAAACACAACAGTGGAATTTCAGAGG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CACAGCTTTAACAGCCCCGGCGTCTTTGTCGTAGAAAACACAACAGTGGAATTTCAGAGG 900

Qy        901 GGCTCCGAGAGGCAAACTTTTAAGATTCCAGGACCTCTGATGGCTGATTTCATCTTCAAG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GGCTCCGAGAGGCAAACTTTTAAGATTCCAGGACCTCTGATGGCTGATTTCATCTTCAAG 960

Qy        961 ACCAGGTACACTGCAGCCAAAGACAGCGTGGTTCAGTTCTTCTTTTACCAGCCCATCAGT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 ACCAGGTACACTGCAGCCAAAGACAGCGTGGTTCAGTTCTTCTTTTACCAGCCCATCAGT 1020

Qy       1021 CATCAGTGGAGACAAACTGACTTCTTTCCCTGCACTGTGACGTGTGGAGGAGGTTATCAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CATCAGTGGAGACAAACTGACTTCTTTCCCTGCACTGTGACGTGTGGAGGAGGTTATCAG 1080
```

```
Qy       1081 CTCAATTCTGCTGAATGTGTGGATATCCGCTTGAAGAGGGTAGTTCCTGACCATTATTGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CTCAATTCTGCTGAATGTGTGGATATCCGCTTGAAGAGGGTAGTTCCTGACCATTATTGT 1140

Qy       1141 CACTACTACCCTGAAAATGTAAAACCAAAACCAAAACTGAAGGAATGCAGCATGGATCCC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CACTACTACCCTGAAAATGTAAAACCAAAACCAAAACTGAAGGAATGCAGCATGGATCCC 1200

Qy       1201 TGCCCATCAAGTGATGGATTTAAAGAGATAATGCCCTATGACCACTTCCAACCTCTTCCT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TGCCCATCAAGTGATGGATTTAAAGAGATAATGCCCTATGACCACTTCCAACCTCTTCCT 1260

Qy       1261 CGCTGGGAACATAATCCTTGGACTGCATGTTCCGTGTCCTGTGGAGGAGGGATTCAGAGA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CGCTGGGAACATAATCCTTGGACTGCATGTTCCGTGTCCTGTGGAGGAGGGATTCAGAGA 1320

Qy       1321 CGGAGCTTTGTGTGTGTAGAGGAATCCATGCATGGAGAGATATTGCAGGTGGAAGAATGG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CGGAGCTTTGTGTGTGTAGAGGAATCCATGCATGGAGAGATATTGCAGGTGGAAGAATGG 1380

Qy       1381 AAGTGCATGTACGCACCCAAACCCAAGGTTATGCAAACTTGTAATCTGTTTGATTGCCCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 AAGTGCATGTACGCACCCAAACCCAAGGTTATGCAAACTTGTAATCTGTTTGATTGCCCC 1440

Qy       1441 AAGTGGATTGCCATGGAGTGGTCTCAGTGCACAGTGACTTGTGGCCGAGGGTTACGGTAC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 AAGTGGATTGCCATGGAGTGGTCTCAGTGCACAGTGACTTGTGGCCGAGGGTTACGGTAC 1500

Qy       1501 CGGGTTGTTCTGTGTATTAACCACCGCGGAGAGCATGTTGGGGGCTGCAATCCACAACTG 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 CGGGTTGTTCTGTGTATTAACCACCGCGGAGAGCATGTTGGGGGCTGCAATCCACAACTG 1560

Qy       1561 AAGTTACACATCAAAGAAGAATGTGTCATTCCCATCCCGTGTTATAAACCAAAAGAAAAA 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 AAGTTACACATCAAAGAAGAATGTGTCATTCCCATCCCGTGTTATAAACCAAAAGAAAAA 1620

Qy       1621 AGTCCAGTGGAAGCAAAATTGCCTTGGCTGAAACAAGCACAAGAACTAGAAGAGACCAGA 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 AGTCCAGTGGAAGCAAAATTGCCTTGGCTGAAACAAGCACAAGAACTAGAAGAGACCAGA 1680

Qy       1681 ATAGCAACAGAAGAACCAACGTTCATTCCAGAACCCTGGTCAGCCTGCAGTACCACGTGT 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 ATAGCAACAGAAGAACCAACGTTCATTCCAGAACCCTGGTCAGCCTGCAGTACCACGTGT 1740

Qy       1741 GGGCCAGGTGTGCAGGTCCGCGAGGTGAAGTGCCGTGTGCTCCTCACATTCACGCAGACT 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 GGGCCAGGTGTGCAGGTCCGCGAGGTGAAGTGCCGTGTGCTCCTCACATTCACGCAGACT 1800

Qy       1801 GAGACTGAGCTGCCCGAGGAAGAGTGTGAAGGCCCCAAGCTGCCCACCGAACGGCCCTGC 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 GAGACTGAGCTGCCCGAGGAAGAGTGTGAAGGCCCCAAGCTGCCCACCGAACGGCCCTGC 1860

Qy       1861 CTCCTGGAAGCATGTGATGAGAGCCCGGCCTCCCGAGAGCTAGACATCCCTCTCCCTGAG 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 CTCCTGGAAGCATGTGATGAGAGCCCGGCCTCCCGAGAGCTAGACATCCCTCTCCCTGAG 1920

Qy       1921 GACAGTGAGACGACTTACGACTGGGAGTACGCTGGGTTCACCCCTTGCACAGCAACATGC 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 GACAGTGAGACGACTTACGACTGGGAGTACGCTGGGTTCACCCCTTGCACAGCAACATGC 1980

Qy       1981 TTGGGAGGCCATCAAGAAGCCATAGCAGTGTGCTTACATATCCAGACCCAGCAGACAGTC 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 TTGGGAGGCCATCAAGAAGCCATAGCAGTGTGCTTACATATCCAGACCCAGCAGACAGTC 2040

Qy       2041 AATGACAGCTTGTGTGATATGGTCCACCGTCCTCCAGCCATGAGCCAGGCCTGTAACACA 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 AATGACAGCTTGTGTGATATGGTCCACCGTCCTCCAGCCATGAGCCAGGCCTGTAACACA 2100

Qy       2101 GAGCCCTGTCCCCCCAGGTGGCATGTGGGCTCTTGGGGGCCCTGCTCAGCTACCTGTGGA 2160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 GAGCCCTGTCCCCCCAGGTGGCATGTGGGCTCTTGGGGGCCCTGCTCAGCTACCTGTGGA 2160

Qy       2161 GTTGGAATTCAGACCGAGATGTGTACTGGCCTGCACCCCAGGGAGACCCCTGCCCCTCCT 2220
```

```
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 GTTGGAATTCAGACCCGAGATGGTACTGCCTGCACCCAGGGGAGACCCCTGCCCACTCCT 2220

Qy       2221 GAGGAGTGCCGAGATGAAAAGCCCCATGCTTTACAAGCATGCAATCAGTTTGACTGCCCT 2280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 GAGGAGTGCCGAGATGAAAAGCCCCATGCTTTACAAGCATGCAATCAGTTTGACTGCCCT 2280

Qy       2281 CCTGGCTGGCACATTGAAGAATGGCAGCAGTGTTCCAGGACTTGTGGGCGGGGAACTCAG 2340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 CCTGGCTGGCACATTGAAGAATGGCAGCAGTGTTCCAGGACTTGTGGGCGGGGAACTCAG 2340

Qy       2341 AACAGAAGAGTCACCTGTGTGCAGCTGCTAACGGATGGCAGCTTTTTGAATCTCTCAGAT 2400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 AACAGAAGAGTCACCTGTGTGCAGCTGCTAACGGATGGCAGCTTTTTGAATCTCTCAGAT 2400

Qy       2401 GAATTGTGCCAAGGACCCAAGGATGCTGCACAAGCCCTGTGCCAGGACAGACTGTGCCCT 2460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 GAATTGTGCCAAGGACCCAAGGATGCTGCACAAGCCCTGTGCCAGGACAGACTGTGCCCT 2460

Qy       2461 CCACATTTAGCTGGTGGGAGACTGGTGAAT 2490
              ||||||||||||||||||||||||||||||
Db       2461 CCACATTTAGCTGGTGGGAGACTGGTGAAT 2490


RESULT 5
US-09-784-358-13
; Sequence 13, Application US/09784358
; Patent No. US20020090027A1
; GENERAL INFORMATION:
;  APPLICANT: Donoho, Gregory
;  APPLICANT:  Scoville, John
;  APPLICANT:  Turner, C. Alexander Jr.
;  APPLICANT:  Friedrich, Glenn
;  APPLICANT:  Zambrowicz, Brian
;  APPLICANT:  Sands, Arthur T.
;  TITLE OF INVENTION: NOVEL HUMAN THROMBOSPONDIN REPEAT PROTEINS AND
;  TITLE OF INVENTION:  POLYNUCLEOTIDES ENCODING THE SAME
;  FILE REFERENCE: LEX 0134 USA
;  CURRENT APPLICATION NUMBER: US/09/784,358
;  CURRENT FILING DATE:  2001-02-15
;  PRIOR APPLICATION NUMBER: US 60/183,282
;  PRIOR FILING DATE: 2000-02-17
;  NUMBER OF SEQ ID NOS: 17
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13
;   LENGTH: 2310
;   TYPE: DNA
;   ORGANISM: homo sapiens
US-09-784-358-13

  Query Match             42.0%;  Score 2180;  DB 10;  Length 2310;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 2180;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        311 CTGGAAGGAATTGTGAAGGCAGAACATTCGGTACAAGACATGCAGCAATCATGACTTCCC 370
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         89 CTGGAAGGAATTGTGAAGGCAGAACATTCGGTACAAGACATGCAGCAATCATGACTTCCC 148

Qy        371 CTCCAGATGCAGAAGATTTCAGAGCCCAGCAGTGCTCAGCCTACAATGATGTCCAGTATC 430
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        149 CTCCAGATGCAGAAGATTTCAGAGCCCAGCAGTGCTCAGCCTACAATGATGTCCAGTATC 208

Qy        431 AGGGGCATTACTATGAATGGCTTCCACGATATAATGATCCTGCTGCCCCGTGTGCACTTA 490
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        209 AGGGGCATTACTATGAATGGCTTCCACGATATAATGATCCTGCTGCCCCGTGTGCACTTA 268

Qy        491 ACTGTCATGCACAAGGACAAAACTTGGTGGTGGAGCTGGCACCTAAGGTACTGGATGGAA 550
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        269 ACTGTCATGCACAAGGACAAAACTTGGTGGTGGAGCTGGCACCTAAGGTACTGGATGGAA 328

Qy        551 CTGGTTGCAACAGGGACTCCTTGGACATGGTGTATCAGTGGCATCTGTCAGGCAGTGGCT 610
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        329 CTGGTTGCAACAGGGACTCCTTGGACATGGTGTATCAGTGGCATCTGTCAGGCAGTGGCT 388

Qy        611 GCGATCGGCAACTGGGAAGCAATGCCAAGGAGGACAACTGTGCAGTCTGTGCCGGCGATG 670
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        389 GCGATCGGCAACTGGGAAGCAATGCCAAGGAGGACAACTGTGCAGTCTGTGCCGGCGATG 448

Qy        671 GCTCCACCTGCAGGCTTGTACGGGGACAATCAAAGTCACACGTTTCTCCTGAAAAAAGAG 730
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        449 GCTCCACCTGCAGGCTTGTACGGGGACAATCAAAGTCACACGTTTCTCCTGAAAAAAGAG 508

Qy        731 AAGAAAATGTAATTGCTGTTCCTTTGGGAAGTCGAAGTGTGAGAATTACAGTGAAAGGAC 790
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        509 AAGAAAATGTAATTGCTGTTCCTTTGGGAAGTCGAAGTGTGAGAATTACAGTGAAAGGAC 568

Qy        791 CTGCCACCTCTTTATTGAATCAAAAAACACTTCAAGGAAGAAAGGAGAACACAGCTTTA 850
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        569 CTGCCACCTCTTTATTGAATCAAAAAACACTTCAAGGAAGAAAGGAGAACACAGCTTTA 628

Qy        851 ACAGCCCGGGGTCTTTGTCCTAGAAAACACAACAGTGGAATTTAGAGGGCCTCCGAGA 910
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        629 ACAGCCCGGGGTCTTTGTCCTAGAAAACACAACAGTGGAATTTAGAGGGCCTCCGAGA 688

Qy        911 GGCAAACTTTTAAGATTCCACGACCTCTGATGGTGATTTCATCTTCAAGACCAGGTACA 970
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        689 GGCAAACTTTTAAGATTCCACGACCTCTGATGGTGATTTCATCTTCAAGACCAGGTACA 748

Qy        971 CTGCAGCCAAAGACAGCGTGGTTCAGTTCTTCTTTACCAGCCCATCAGTCATCAGTGGA 1030
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        749 CTGCAGCCAAAGACAGCGTGGTTCAGTTCTTCTTTACCAGCCCATCAGTCATCAGTGGA 808

Qy       1031 GACAAACTGACTTCTTTCCCTGCACTGTGACGGTGGAGGAGGTTATCAGCTCAATTCTG 1090
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        809 GACAAACTGACTTCTTTCCCTGCACTGTGACGGTGGAGGAGGTTATCAGCTCAATTCTG 868

Qy       1091 CTGAATGTGTGGATATCCGCTTGAAGAGGGTAGTTCCTGACCATTATTGTCACTACTACC 1150
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        869 CTGAATGTGTGGATATCCGCTTGAAGAGGGTAGTTCCTGACCATTATTGTCACTACTACC 928

Qy       1151 CTGAAAATGTAAAACCAAAACCAAAACTGAAGGAATGCAGCATGGATCCCTGCCCATCAA 1210
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        929 CTGAAAATGTAAAACCAAAACCAAAACTGAAGGAATGCAGCATGGATCCCTGCCCATCAA 988

Qy       1211 GTGATGGATTTAAAGAGATAATGCCCTATGACCACTTCCAACCTCTTCCTCGCTGGGAAC 1270
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        989 GTGATGGATTTAAAGAGATAATGCCCTATGACCACTTCCAACCTCTTCCTCGCTGGGAAC 1048

Qy       1271 ATAATCCTTGGACTGCATGTTCCGTGTCCTGTGGAGGAGGGATTCAGAGACGGAGCTTTG 1330
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1049 ATAATCCTTGGACTGCATGTTCCGTGTCCTGTGGAGGAGGGATTCAGAGACGGAGCTTTG 1108

Qy       1331 TGTGTGTAGAGGAATCCATGCATGGAGAGATATTGCAGGTGGAAGAATGGAAGTGCATGT 1390
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1109 TGTGTGTAGAGGAATCCATGCATGGAGAGATATTGCAGGTGGAAGAATGGAAGTGCATGT 1168

Qy       1391 ACGCACCCAAACCCAAGGTTATGCAAACTTGTAATCTGTTTGATTGCCCCAAGTGGATTG 1450
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1169 ACGCACCCAAACCCAAGGTTATGCAAACTTGTAATCTGTTTGATTGCCCCAAGTGGATTG 1228

Qy       1451 CCATGGAGTGGTCTCAGTGCACAGTGACTTGTGGCCGAGGGTTACGGTACCGGGTGGTTC 1510
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1229 CCATGGAGTGGTCTCAGTGCACAGTGACTTGTGGCCGAGGGTTACGGTACCGGGTGGTTC 1288

Qy       1511 TGTGTATTAACCACCGCGGAGAGCATGTTGGGGGCTGCAATCCACAACTGAAGTTACACA 1570
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1289 TGTGTATTAACCACCGCGGAGAGCATGTTGGGGGCTGCAATCCACAACTGAAGTTACACA 1348

Qy       1571 TCAAAGAAGAATGTGTCATTCCCATCCGTGTTTATAAACCAAAAGAAAAAAGTCCAGTGG 1630
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1349 TCAAAGAAGAATGTGTCATTCCCATCCGTGTTTATAAACCAAAAGAAAAAAGTCCAGTGG 1408

Qy       1631 AAGCAAAATTGCCTTGGCTGAAACAAGCACAAGAACTAGAAGAGACCAGAATAGCAACAG 1690
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1409 AAGCAAAATTGCCTTGGCTGAAACAAGCACAAGAACTAGAAGAGACCAGAATAGCAACAG 1468
```

```
Qy       1691 AAGAACCAAGCTTCATTCCAGAACCCTGCTCAGCCTCCAGTACCACCTCTGGCCCAGCTC 1750
              |||||||||| ||||||||||||||||||||||||| ||||||||| | |||||||||||
Db       1469 AAGAACCAAGGTTCATTCCAGAACCCTGCTCAGCCTGCAGTACCACGTGTGGCCCAGCTC 1528

Qy       1751 TGCAGGTCCGCGAGGTGAAGTGCCGTGTGCTCCTCACATTCACGCAGACTGAGACTGAGC 1810
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1529 TGCAGGTCCGCGAGGTGAAGTGCCGTGTGCTCCTCACATTCACGCAGACTGAGACTGAGC 1588

Qy       1811 TGCCGGAGGAAGAGTGTGAAGGCCCCAAGCTGCCCACCGAACGGCCCTGCCTCCTGGAAG 1870
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1589 TGCCGGAGGAAGAGTGTGAAGGCCCCAAGCTGCCCACCGAACGGCCCTGCCTCCTGGAAG 1648

Qy       1871 CATGTGATGAGAGCCCGGCCTCCCGAGAGCTAGACATCCCTCTCCCTGAGGACAGTGAGA 1930
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1649 CATGTGATGAGAGCCCGGCCTCCCGAGAGCTAGACATCCCTCTCCCTGAGGACAGTGAGA 1708

Qy       1931 CGACTTACGACTGGGAGTACGCTGGGTTCACCCCTTGCACAGCAACATGCTTGGGAGGCC 1990
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1709 CGACTTACGACTGGGAGTACGCTGGGTTCACCCCTTGCACAGCAACATGCTTGGGAGGCC 1768

Qy       1991 ATCAAGAAGCCATAGCAGTGTGCTTACATATCCAGACCCAGCAGACAGTCAATGACAGCT 2050
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1769 ATCAAGAAGCCATAGCAGTGTGCTTACATATCCAGACCCAGCAGACAGTCAATGACAGCT 1828

Qy       2051 TGTGTGATATGGTCCACCGTCCTCCAGCCATGAGCCAGGCCTGTAACACAGAGCCCTGTC 2110
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1829 TGTGTGATATGGTCCACCGTCCTCCAGCCATGAGCCAGGCCTGTAACACAGAGCCCTGTC 1888

Qy       2111 CCCCCAGGTGGCATGTGGGCTCTTGGGGGCCCTGCTCAGCTACCTGTGGAGTTGCAATTC 2170
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1889 CCCCCAGGTGGCATGTGGGCTCTTGGGGGCCCTGCTCAGCTACCTGTGGAGTTGCAATTC 1948

Qy       2171 AGACCCGAGATGTGTACTGCCTGCACCCAGGGGAGACCCCTGCCCCTCCTGAGGAGTGCC 2230
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1949 AGACCCGAGATGTGTACTGCCTGCACCCAGGGGAGACCCCTGCCCCTCCTGAGGAGTGCC 2008

Qy       2231 GAGATGAAAAGCCCCATGCTTTACAAGCATGCAATCACTTTGACTGCCCTCCTGCCTGGC 2290
              ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db       2009 GAGATGAAAAGCCCCATGCTTTACAAGCATGCAATCAGTTTGACTGCCCTCCTGGCTGGC 2068

Qy       2291 ACATTGAAGAATGGCAGCAGTGTTCCAGGACTTGTGGCGGGGGAACTCAGAACAGAAGAG 2350
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2069 ACATTGAAGAATGGCAGCAGTGTTCCAGGACTTGTGGCGGGGGAACTCAGAACAGAAGAG 2128

Qy       2351 TCACCTGTCGGCAGCTGCTAACGGATGGCAGCTTTTTGAATCTCTCAGATGAATTGTGCC 2410
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2129 TCACCTGTCGGCAGCTGCTAACGGATGGCAGCTTTTTGAATCTCTCAGATGAATTGTGCC 2188

Qy       2411 AAGGACCCAAGGCATCGTCTCACAAGTCCTGTGCCAGGACAGACTGTCCTCCACATTTAG 2470
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2189 AAGGACCCAAGGCATCGTCTCACAAGTCCTGTGCCAGGACAGACTGTCCTCCACATTTAG 2248

Qy       2471 CTGTGGGAGACTGGTCGAAG 2490
              |||||||||||||||||||
Db       2249 CTGTGGGAGACTGGTCGAAT 2268


RESULT 6
US-09-784-358-7
; Sequence 7, Application US/09784358
; Patent No. US20020099027A1
; GENERAL INFORMATION:
;  APPLICANT: Donoho, Gregory
;  APPLICANT:  Scoville, John
;  APPLICANT:  Turner, C. Alexander Jr.
;  APPLICANT:  Friedrich, Glenn
;  APPLICANT:  Zambrowicz, Brian
;  APPLICANT:  Sands, Arthur T.
;  TITLE OF INVENTION: NOVEL HUMAN THROMBOSPONDIN REPEAT PROTEINS AND
;  TITLE OF INVENTION:  POLYNUCLEOTIDES ENCODING THE SAME
;  FILE REFERENCE: LEX-0134-USA
;  CURRENT APPLICATION NUMBER: US/09/784,358
;  CURRENT FILING DATE:  2001-02-15
;  PRIOR APPLICATION NUMBER: US 60/183,282
;  PRIOR FILING DATE: 2000-02-17
;  NUMBER OF SEQ ID NOS: 17
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7
;   LENGTH: 2175
;   TYPE: DNA
;   ORGANISM: homo sapiens
US-09-784-358-7

  Query Match             41.7%;  Score 2119;  DB 10;  Length 2175;
  Best Local Similarity   99.5%;  Pred. No. 0;
  Matches 2125;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          1 ATGGCTTCCTGGACGAGCCCCTGGTGGGTGCTGATAGGGATGGTCTTCATGCACTCTCCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGCTTCCTGGACGAGCCCCTGGTGGGTGCTGATAGGGATGGTCTTCATGCACTCTCCC 60

Qy         61 CTCCCGCAGACCACAGCTGAGAAATCTCCTGGAGCCTATTTCCTTCCCGAGTTTGCACTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTCCCGCAGACCACAGCTGAGAAATCTCCTGGAGCCTATTTCCTTCCCGAGTTTGCACTT 120

Qy        121 TCTCCTCAGGGAAGTTTTCTGGAAGACACAACAGGGGAGCAGTTCCTCACTTATCGCTAT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCTCCTCAGGGAAGTTTTCTGGAAGACACAACAGGGGAGCAGTTCCTCACTTATCGCTAT 180

Qy        181 GATGACCAGACCTCAAGAAACACTCGTTCAGATGAAGACAAAGATGGCAACTGGGATGCT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GATGACCAGACCTCAAGAAACACTCGTTCAGATGAAGACAAAGATGGCAACTGGGATGCT 240

Qy        241 TGGGGCCACTGGACTGACTGCTCCCGGACCTGTGGGGGAGGAGCATCATATTCTCTGCGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGGGGCGACTGGAGTGACTGCTCCCGGACCTGTGGGGGAGGAGCATCATATTCTCTGCGG 300

Qy        301 ACATGTTTGACTGGAAGGAATTGTGAAGGGCAGAACATTCGGTACAAGACATGCAGCAAT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ACATGTTTGACTGGAAGGAATTGTGAAGGGCAGAACATTCGGTACAAGACATGCAGCAAT 360

Qy        361 CATGACTGCCCTCCAGATGCAGAAGATTTCAGAGCCCAGCAGTGCTCAGCCTACAATGAT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CATGACTGCCCTCCAGATGCAGAAGATTTCAGAGCCCAGCAGTGCTCAGCCTACAATGAT 420

Qy        421 GTCCAGTATCAGGGGCATTACTATGAATGGCTTCCACGATATAATGATCCTGCTGCCCCG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GTCCAGTATCAGGGGCATTACTATGAATGGCTTCCACGATATAATGATCCTGCTGCCCCG 480

Qy        481 TGTGCACTCAAGTGTCATGCACAAGGACAAAACTTGGTGGTGGAGCTGGCACCTAAGGTA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TGTGCACTCAAGTGTCATGCACAAGGACAAAACTTGGTGGTGGAGCTGGCACCTAAGGTA 540

Qy        541 CTGGATGGAACTCGTTGCAACACGGACTCCTTGGACATGTGTATCAGTGGCATCTGTCAG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTGGATGGAACTCGTTGCAACACGGACTCCTTGGACATGTGTATCAGTGGCATCTGTCAG 600

Qy        601 GCAGTGGGCTGCGATCGGCAACTGGGAAGCAATGCCAAGGAGGACAACTGTGGAGTCTGT 660
              |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db        601 GCAGTGGGCTGCGATCGGCAACTGGGAAGCAATGCCAAGGAGGACAACTGTGGAGTCTGT 660

Qy        661 GCCGGCGATGGCTCCACCTGCAGGCTTGTACGGGGACAATCAAAGTCACACGTTTCTCCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GCCGGCGATGGCTCCACCTGCAGGCTTGTACGGGGACAATCAAAGTCACACGTTTCTCCT 720

Qy        721 GAAAAAAGAGAAGAAAATGTAATTGCTGTTCCTTTGGGAAGTCGAAGTGTGAGAATTACA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GAAAAAAGAGAAGAAAATGTAATTGCTGTTCCTTTGGGAAGTCGAAGTGTGAGAATTACA 780

Qy        781 GTGAAAGGACCTGCCCACCTCTTTATTGAATCAAAAACACTTCAAGGAAGCAAAGGAGAA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GTGAAAGGACCTGCCCACCTCTTTATTGAATCAAAAACACTTCAAGGAAGCAAAGGAGAA 840

Qy        841 CACAGCTTTAACAGCCCCGGCGTCTTTGTCGTAGAAAACACAACAGTGGAATTTCAGAGG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CACAGCTTTAACAGCCCCGGCGTCTTTGTCGTAGAAAACACAACAGTGGAATTTCAGAGG 900
```

```
Qy        901 GGCTCCGAGAGGCAAACTTTAAGATTCAGGACCTCTGATGGCTGATTTCATCTTCAAG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GGCTCCGAGAGGCAAACTTTAAGATTCAGGACCTCTGATGGCTGATTTCATCTTCAAG 960

Qy        961 ACCAGGTACACTGCAGCCAAAGACAGCGTGGTTCAGTTCTTCTTTACCAGCCCATCAGT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 ACCAGGTACACTGCAGCCAAAGACAGCGTGGTTCAGTTCTTCTTTACCAGCCCATCAGT 1020

Qy       1021 CATCAGGGAGACAAACTGACTTCTTTCCTGCACTGTGACGTGGGAGGAGGTATCAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CATCAGGGAGACAAACTGACTTCTTTCCTGCACTGTGACGTGGGAGGAGGTATCAG 1080

Qy       1081 CTCAATTCTGCTGAATGTGTGGATATCCCTTGAAGAGGGTACTTCCTGACCATTATTGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CTCAATTCTGCTGAATGTGTGGATATCCCTTGAAGAGGGTACTTCCTGACCATTATTGT 1140

Qy       1141 CATTACTACCCTGAAAATTTAAAACTAAAACTAAAACTGAAGGAATGAGCATGGATCCT 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CATTACTACCCTGAAAATTTAAAACTAAAACTAAAACTGAAGGAATGAGCATGGATCCT 1200

Qy       1201 TGCCCATCAAGTGATGGATTTAAAGAGATAATGCCCTATGACCACTTCCAACCTCTTCCT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TGCCCATCAAGTGATGGATTTAAAGAGATAATGCCCTATGACCACTTCCAACCTCTTCCT 1260

Qy       1261 CGCTGGGAACATAATCCTTGGACTGCATGTTCCGTGTCCTGTGGAGGAGGCATTCAGAGA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CGCTGGGAACATAATCCTTGGACTGCATGTTCCGTGTCCTGTGGAGGAGGCATTCAGAGA 1320

Qy       1321 CGGAGCTTTGTGTGTGTAGAGGAATCCATGCATGGAGAGATATTGCAGGTGGAAGAATGG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CGGAGCTTTGTGTGTGTAGAGGAATCCATGCATGGAGAGATATTGCAGGTGGAAGAATGG 1380

Qy       1381 AAGTGCATGTACGCACCCAAACCCAAGGTTATGCAAACTTGTAATCTGTTTGATTGCCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 AAGTGCATGTACGCACCCAAACCCAAGGTTATGCAAACTTGTAATCTGTTTGATTGCCC 1440

Qy       1441 AAGTGGATTGCCATGGATTGCCCTCAGTGCACAGTGACTTGTGGCCGAGGGTTACGGTAC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 AAGTGGATTGCCATGGATTGCCCTCAGTGCACAGTGACTTGTGGCCGAGGGTTACGGTAC 1500

Qy       1501 CGGGTTGTTCTGTGTATTAACCACCGGGGACACCATGCT[illegible] 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 CGGGTTGTTCTGTGTATTA[illegible]

Qy       1561 AAGTTATATATTAAATAAGAAGGTGTATCCCATCCCGTGTTATAAACTAAAAGAAAAA 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 AAGTTATATATTAAATAAGAAGGTGTATCCCATCCCGTGTTATAAACTAAAAGAAAAA 1620

Qy       1621 AGTCCATTGGAATTAAAATTCCTTGCTTCAAAAAACACAAGAATAGAAGAGACTAGA 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 AGTCCATTGGAATTAAAATTCCTTGCTTCAAAAAACACAAGAATAGAAGAGACTAGA 1680

Qy       1681 ATAGCAACAGAAGAACCAACGTTCATTCCAGAACCTGGTCAGCCTGCAGTTACCACGTGT 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 ATAGCAACAGAAGAACCAACGTTCATTCCAGAACCTGGTCAGCCTGCAGTTACCACGTGT 1740

Qy       1741 GGGCCAGGTGTGCAGGTCCGCTAGGTGAAGGCCGTGTGCTCCTCACATTCACGCAGACT 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 [illegible]

Qy       1801 GAGATGGAGCTGCCCGAGGAAGAGTGTGAAGGCCCCAAGCTGCCCACCGAACCCCCCTGCT 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 GAGATGGAGCTGCCCGAGGAAGAGTGTGAAGGCCCCAAGCTGCCCACCGAACCCCCCTGCT 1860

Qy       1861 CTCCTGGAAGCATCTGATGAGAGCCCCGGCTCCCGAGAGCTAGACATCCCTGCCCTGAG 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 CTCCTGGAAGCATCTGATGAGAGCCCCGGCTCCCGAGAGCTAGACATCCCTGCCCTGAG 1920

Qy       1921 GACATGAGACGACTTACGACTGGGAGTTACCCTGGTGACCCTTGTACAGCAACATGCC 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 GACATGAGACGACTTACGACTGGGAGTTACCCTGGTGACCCTTGTACAGCAACATGCC 1980

Qy       1981 TTGGGAGGCCATCAAGAAGCCATAGCAGTGTGCTTACATATCCAGACCCAGCAGACAGTC 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 TTGGGAGGCCATCAAGAAGCCATAGCAGTGTGCTTACATATCCAGACCCAGCAGACAGTC 2040

Qy       2041 AATGACAGCTTGTGTGATATGGTCCACCGTCCTCCAGCCATGAGCCAGGCCTGTAACACA 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 AATGACAGCTTGTGTGATATGGTCCACCGTCCTCCAGCCATGAGCCAGGCCTGTAACACA 2100

Qy       2101 GAGCCCTGTCCCCCAGGTTCCATGTGGGCTCTTG 2134
              ||||||||||||||||  |    |   |||
Db       2101 GAGCCCTGTCCCCCAGGAGAGAGCCAGCAGCTTG 2134


RESULT 7
US-09-784-358-9
; Sequence 9, Application US/09784358
; Patent No. US20020099027A1
; GENERAL INFORMATION:
;  APPLICANT: Donoho, Gregory
;  APPLICANT:  Scoville, John
;  APPLICANT:  Turner, C. Alexander Jr.
;  APPLICANT:  Friedrich, Glenn
;  APPLICANT:  Zambrowicz, Brian
;  APPLICANT:  Sands, Arthur T.
;  TITLE OF INVENTION: NOVEL HUMAN THROMBOSPONDIN REPEAT PROTEINS AND
;  TITLE OF INVENTION:  POLYNUCLEOTIDES ENCODING THE SAME
;  FILE REFERENCE: LEX-0134-USA
;  CURRENT APPLICATION NUMBER: US/09/784,358
;  CURRENT FILING DATE:  2001-02-15
;  PRIOR APPLICATION NUMBER: US 60/183,282
;  PRIOR FILING DATE: 2000-02-17
;  NUMBER OF SEQ ID NOS: 17
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9
;   LENGTH: 1953
;   TYPE: DNA
;   ORGANISM: homo sapiens
US-09-784-358-9

  Query Match             85.0%;  Score 1815;  DB 10;  Length 1953;
  Best Local Similarity   99.5%;  Pred. No. 0;
  Matches 1815;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy        311 [illegible] 370
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         89 CCTCTGGAAGGAATTGTGAAGCCAGAACATTCGGTACAAGACATCACCAATCATGACTGCC 148

Qy        371 CCTCAGATGCAGAAGATTTCAGGCCAGCAGCCTCACCTCACAAATGATGCCAGTATC 430
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        149 [illegible] 208

Qy        431 AGGGGCATTACTATGAATGGCTTCCACGATATAATGATCCGCTGCCCCGTGTGCACTCA 490
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        209 AGGGGCATTACTATGAATGGCTTCCACGATATAATGATCCGCTGCCCCGTGTGCACTCA 268

Qy        491 AGTGTCATGCACAAGGACAAAACTTGGTGGTGGAGCTGGCACCTAAGGTACTGGATGGAA 550
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        269 AGTGTCATGCACAAGGACAAAACTTGGTGGTGGAGCTGGCACCTAAGGTACTGGATGGAA 328

Qy        551 CTCCGTTGCAACACTGACTCCTTGGACATGGTTATCAGTCCCATCTCTCAGGCAGTGGGCT 610
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        329 CTCCGTTGCAACACTGACTCCTTGGACATGGTTATCAGTCCCATCTCTCAGGCAGTGGGCT 388

Qy        611 GCGATCGGCAACTGGAAGCAATGCCAAGGACTACAACTTGGAGTCTGGCCTGGTGATG 670
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        389 GCGATCGGCAACTGGAAGCAATGCCAAGGACTACAACTTGGAGTCTGGCCTGGTGATG 448

Qy        671 GCTCCACCTGCAGGCTTGTACGGGGACAATCAAAGTCACACGTCCTCTTGAAAAAAGAG 730
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        449 GCTCCACCTGCAGGCTTGTACGGGGACAATCAAAGTCACACGTCCTCTTGAAAAAAGAG 508

Qy        731 AAGAAAATGTAATTGCTGTTCCTTTGGGAAGTGAAGTGAGAATTACACTGAAAGGAC 790
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        509 AAGAAAATGTAATTGCTGTTCCTTTGGGAAGTCGAAGTGTGAGAATTACAGTGAAAGGAC 568

Qy        791 CTGCCCACCTCTTTATTGAATCAAAAACACTTCAAGGAAGCAAAGGAGAACACAGCTTTA 850
              |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db        569 CTGCCCACCTCTTTATTGAATCAAAAACACTTCAAGGAAGCAAAGGAGAACACAGCTTTA 628

Qy        851 ACAGCCCCGGCGTCTTTGTCGTAGAAAACACAACAGTGGAATTTCAGAGGGGCTCCGAGA 910
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        629 ACAGCCCCGGCGTCTTTGTCGTAGAAAACACAACAGTGGAATTTCAGAGGGGCTCCGAGA 688

Qy        911 GGCAAACTTTTAAGATTCCAGGACCTCTGATGGCTGATTTCATCTTCAAGACCAGGTACA 970
              |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db        689 GGCAAACTTTTAAGATTCCAGGACCTCTGATGGCTGATTTCATCTTCAAGACCAGGTACA 748

Qy        971 CTGCAGCCAAAGACAGCGTGGTTCAGTTCTTCTTTTACCAGCCCATCAGTCATCAGTGGA 1030
              |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db        749 CTGCAGCCAAAGACAGCGTGGTTCAGTTCTTCTTTTACCAGCCCATCAGTCATCAGTGGA 808

Qy       1031 GACAAACTGACTTCTTTCCCTGCACTGTGACGTGTGGAGGAGGTTATCAGCTCAATTCTG 1090
              |||| ||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db        809 GACAAACTGACTTCTTTCCCTGCACTGTGACGTGTGGAGGAGGTTATCAGCTCAATTCTG 868

Qy       1091 CTGAATGTGTGGATATCCGCTTGAAGAGGGTAGTTCCTGACCATTATTGTCACTACTACC 1150
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        869 CTGAATGTGTGGATATCCGCTTGAAGAGGGTAGTTCCTGACCATTATTGTCACTACTACC 928

Qy       1151 CTGAAAATGTAAAACCAAAACCAAAACTGAAGGAATGCAGCATGGATCCCTGCCCATCAA 1210
              ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        929 CTGAAAATGTAAAACCAAAACCAAAACTGAAGGAATGCAGCATGGATCCCTGCCCATCAA 988

Qy       1211 GTGATGGATTTAAAGAGATAATGCCCTATGACCACTTCCAACCTCTTCCTCGCTGGGAAC 1270
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        989 GTGATGGATTTAAAGAGATAATGCCCTATGACCACTTCCAACCTCTTCCTCGCTGGGAAC 1048

Qy       1271 ATAATCCTTGGACTGCATGTTCCGTGTCCTGTGGACCACCCATTCACACACCGACCTTTG 1330
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1049 ATAATCCTTGGACTGCATGTTCCGTGTCCTGTGGAGGAGGCATTCACACACCCAGCTTTG 1108

Qy       1331 TGTGTGTAGAGGAATCCATGCATGGAGAGATATTGCACGTGGAACAATGGAAGTGCATGT 1390
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1109 TGTGTGTAGAGGAATCCATGCATGGAGAGATATTGCAGGTGGAAGAATGGAAGTGCATGT 1168

Qy       1391 ACGCACCCAAACCCAAGGTTATGCAAACTTGTAATCTGTTTGATTGCCCCAAGTGGATTG 1450
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1169 ACGCACCCAAACCCAAGGTTATGCAAACTTGTAATCTGTTTGATTGCCCCAAGTGGATTG 1228

Qy       1451 CCATGGAGTGGTCTCAGTGCACAGTGACTTGTGGCCGAGGGTTACGGTACCGGGTTGTTC 1510
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1229 CCATGGAGTGGTCTCAGTGCACAGTGACTTGTGGCCGAGGGTTACGGTACCGGGTTGTTC 1288

Qy       1511 TGTGTATTAACCACCGCGGAGAGCATGTTGGGGGCTGCAATCCACAACTGAAGTTACACA 1570
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1289 TGTGTATTAACCACCGCGGAGAGCATGTTGGGGGCTGCAATCCACAACTGAAGTTACACA 1348

Qy       1571 TCAAAGAAGAATGTGTCATTCCCATCCCGTGTTATAAACCAAAAGAAAAAAGTCCAGTGG 1630
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1349 TCAAAGAAGAATGTGTCATTCCCATCCCGTGTTATAAACCAAAAGAAAAAAGTCCAGTGG 1408

Qy       1631 AAGCAAAATTGCCTTGGCTGAAACAAGCACAAGAACTAGAAGAGACCAGAATAGCAACAG 1690
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1409 AAGCAAAATTGCCTTGGCTGAAACAAGCACAAGAACTAGAAGAGACCAGAATAGCAACAG 1468

Qy       1691 AAGAACCAACGTTCATTCCAGAACCCTGGTCAGCCTGCAGTACCACGTGTGGGCCAGGTG 1750
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1469 AAGAACCAACGTTCATTCCAGAACCCTGGTCAGCCTGCAGTACCACGTGTGGGCCAGGTG 1528

Qy       1751 TGCAGGTCCGCGAGGTGAAGTGCCGTGTGCTCCTCACATTCACGCAGACTGAGACTGAGC 1810
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1529 TGCAGGTCCGCGAGGTGAAGTGCCGTGTGCTCCTCACATTCACGCAGACTGAGACTGAGC 1588

Qy       1811 TGCCCGAGGAAGAGTGTGAAGGCCCCAAGCTGCCCACCGAACGGCCCTGCCTCCTGGAAG 1870
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1589 TGCCCGAGGAAGAGTGTGAAGGCCCCAAGCTGCCCACCGAACGGCCCTGCCTCCTGGAAG 1648

Qy       1871 CATGTGATGAGAGCCCGGCCTCCCGAGAGCTAGACATCCCTCTCCCTGAGGACAGTGAGA 1930
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1649 CATGTGATGAGAGCCCGGCCTCCCGAGAGCTAGACATCCCTCTCCCTGAGGACAGTGAGA 1708

Qy       1931 CGACTTACGACTGGGAGTACGCTGGGTTCACCCCTTGCACAGCAACATGCTTGGGAGGCC 1990
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1709 CGACTTACGACTGGGAGTACGCTGGGTTCACCCCTTGCACAGCAACATGCTTGGGAGGCC 1768

Qy       1991 ATCAAGAAGCCATAGCAGTGTGCTTACATATCCAGACCCAGCAGACAGTCAATGACAGCT 2050
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1769 ATCAAGAAGCCATAGCAGTGTGCTTACATATCCAGACCCAGCAGACAGTCAATGACAGCT 1828

Qy       2051 TGTGTGATATGGTCCACCGTCCTCCAGCCATGAGCCAGGCCTGTAACACAGAGCCCTGTC 2110
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1829 TGTGTGATATGGTCCACCGTCCTCCAGCCATGAGCCAGGCCTGTAACACAGAGCCCTGTC 1888

Qy       2111 CCCCCAGGTGGCATGTGGGCTCTTG 2135
              |||||||| |  |    |   ||||
Db       1889 CCCCCAGGAGAGAGCCAGCAGCTTG 1913


RESULT 8
US-09-784-358-3
; Sequence 3, Application US/09784358
; Patent No. US20020099027A1
; GENERAL INFORMATION:
;  APPLICANT: Donoho, Gregory
;  APPLICANT:  Scoville, John
;  APPLICANT:  Turner, C. Alexander Jr.
;  APPLICANT:  Friedrich, Glenn
;  APPLICANT:  Zambrowicz, Brian
;  APPLICANT:  Sands, Arthur T.
;  TITLE OF INVENTION: NOVEL HUMAN THROMBOSPONDIN REPEAT PROTEINS AND
;  TITLE OF INVENTION:  POLYNUCLEOTIDES ENCODING THE SAME
;  FILE REFERENCE: LEX-0134-USA
;  CURRENT APPLICATION NUMBER: US/09/784,358
;  CURRENT FILING DATE:  2001-02-15
;  PRIOR APPLICATION NUMBER: US 60/183,282
;  PRIOR FILING DATE: 2000-02-17
;  NUMBER OF SEQ ID NOS: 17
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 1341
;   TYPE: DNA
;   ORGANISM: homo sapiens
US-09-784-358-3

  Query Match             26.1%;  Score 1327;  DB 10;  Length 1341;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1339;  Conservative    0;  Mismatches    0;  Indels    2;  Gaps    1;

Qy          1 ATGGCTTCCTGGACGAGCCCCTGGTGGTGCTGATAGGGATGGTCTTCATGCACTCTCCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGCTTCCTGGACGAGCCCCTGGTGGTGCTGATAGGGATGGTCTTCATGCACTCTCCC 60

Qy         61 CTCCCGCAGACCACAGCTGAGAAATCTCCTGGAGCCTATTTCCTTCCCGAGTTTGCACTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTCCCGCAGACCACAGCTGAGAAATCTCCTGGAGCCTATTTCCTTCCCGAGTTTGCACTT 120

Qy        121 TCTCCTCAGGGAAGTTTTCTGGAAGACACAACAGGGGAGCAGTTCCTCACTTATCGCTAT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCTCCTCAGGGAAGTTTTCTGGAAGACACAACAGGGGAGCAGTTCCTCACTTATCGCTAT 180

Qy        181 GATGACCAGACCTCAAGAAACACTCGTTCAGATGAAGACAAAGATGGCAACTGGGATGCT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GATGACCAGACCTCAAGAAACACTCGTTCAGATGAAGACAAAGATGGCAACTGGGATGCT 240

Qy        241 TGGGGCGACTGGAGTGACTGCTCCCGGACCTGTGGGGGAGGAGCATCATATTCTCTGCGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGGGGCGACTGGAGTGACTGCTCCCGGACCTGTGGGGGAGGAGCATCATATTCTCTGCGG 300

Qy        301 AGATGTTTGACTGGAAGGAATTGTGAAGGGCAGAACATTCGGTACAAGACATGCAGCAAT 360
```

```
Db        301 AGATGTTGACTGGAAGGAATTGTGAAGGCAGAACATTCGGTACAAGACATGCAGCAAT 360

Qy        361 CATGACTGCCCTCCAGATGCAGAAGATTTCAGAGCCAGCAGTGCTCAGCCTACAATGAT 420
Db        361 CATGACTGCCCTCCAGATGCAGAAGATTTCAGAGCCAGCAGTGCTCAGCCTACAATGAT 420

Qy        421 GTCCATTATCAGGGGATTACTATGAATGGCTTCCACGATATAATGAT[illegible] 480
Db        421 GTCCATTATCAGGGGATTACTATGAATGGCTTCCACGATATAATGAT[illegible] 480

Qy        481 TGTCCACTCAAGTGTCATGCA AAGGACAAAACTTGG[illegible]ACTTAAGGTA 540
Db        481 TGTCCACTCAAGTGTCATGCATAAGGACAAAACTTGG[illegible]ACTTAAGGTA 540

Qy        541 CTGGATGGAACTGGTTGCAACACGGACTCCTTGGACATGG[illegible]ATCTGTCAT 600
Db        541 CTGGATGGAACTGGTTGCAACACGGACTCCTTGGACATGGTGTATCAGTGGATCTGTCAG 600

Qy        601 GCAGTGGGCTGGGATCGGCAACTGGGAAGCAATGGCAAGGAGGACAACTGTGGAGTCTGT 660
Db        601 GCAGTGGGCTGGGATCGGCAACTGGGAAGCAATGGCAAGGAGGACAACTGTGGAGTCTGT 660

Qy        661 GCCGGCGATGGCTCCACCTGCAGGCTTGTACGGGGACAATCAAAGTCACAGCTTTCTCCT 720
Db        661 GCCGGCGATGGCTCCACCTGCAGGCTTGTACGGGGACAATCAAAGTCACAGCTTTCTCCT 720

Qy        721 GAAAAAAGAGAAGAAAATGTAATTGCTGTTCCTTGGGAAGTGGAAGTGTGAGAATTACA 780
Db        721 GAAAAAAGAGAAGAAAATGTAATTGCTGTTCCTTGGGAAGTGGAAGTGTGAGAATTACA 780

Qy        781 GTGAAAGGACCTGCCCACCTCTTTATTGAATCAAAAACACTTCAAGGAAGCAAAGGAGAA 840
Db        781 GTGAAAGGACCTGCCCACCTCTTTATTGAATCAAAAACACTTCAAGGAAGCAAAGGAGAA 840

Qy        841 CACAGCTTTAACAGCCCCGGCGTCTTTGTCGTAGAAAACACAACAGTGGAATTTCAGAGG 900
Db        841 CACAGCTTTAACAGCCCCGGCGTCTTTGTCGTAGAAAACACAACAGTGGAATTTCAGAGG 900

Qy        901 GGCTCCGAGAGGCAAACTTTTAAGATTCCAGGACCTCTGATGGCTGATTTCATCTTCAAT 960
Db        901 GGCTCCGAGAGGCAAACTTTTAAGATTCCAGGACCTCTGATGGCTGATTTCATCTTCAAT 960

Qy        961 ACCAGGTACACTGGAGCCAAAGACAGGCTGGTTCAGTTCTTCTTTACCAGCCCATCACT 1020
Db        961 ACCAGGTACACTGGAGCCAAAGACAGGCTGGTTCAGTTCTTCTTTACCAGCCCATCACT 1020

Qy       1021 CATCAGTGGAGACAAACTGACTTCTTTCCCTGCACTGTGACGTGTGGAGGAGGTTATCAG 1080
Db       1021 CATCAGTGGAGACAAACTGACTTCTTTCCCTGCACTGTGACGTGTGGAGGAGGTTATCAG 1080

Qy       1081 CTCAATTCTGCTGAATGTGTGGATATCCGCTTGAAGAGGGTAGTTCCTGACCATTATTGT 1140
Db       1081 CTCAATTCTGCTGAATGTGTGGATATCCGCTTGAAGAGGGTAGTTCCTGACCATTATTGT 1140

Qy       1141 CACTACTACCCTGAAAATGTAAAACCAAAACCAAAACTGAAGGAATGCAGCATGGATCCC 1200
Db       1141 CACTACTACCCTGAAAATGTAAAACCAAAACCAAAACTGAAGGAATGCAGCATGGATCCC 1200

Qy       1201 TGCCCATCAAGTGATGGATTTAAAGAGATAATGCCCTATGACCACTTCCAACCTCTTCCT 1260
Db       1201 TGCCCATCAAGTGATGGATTTAAAGAGATAATGCCCTATGACCACTTCCAACCTCTTCCT 1260

Qy       1261 CGAGCTGGGAACATAATCCTTGGACTGCATGTTCCGTGTCCTGTGGAGGAGGGATTCAGA 1318
Db       1261 CGAGCTGGGAACATAATCCTTGGACTGCATGTTCCGTGTCCTGTGGAGGAGGGATTCAGA 1320

Qy       1319 GACGGAGCTTTGTGTGTGTAT 1339
Db       1321 GACGGAGCTTTGTGTGTGTAT 1341
```

RESULT 9

```
US-10-044-807-1
; Sequence 1, Application US/10044807
; Patent No. US20020165187A1
; GENERAL INFORMATION:
;  APPLICANT: Yu, Xuanchuan
;  APPLICANT: Turner, C. Alexander Jr.
;  TITLE OF INVENTION: No. US20020165187A1e: Human Protease and Polynucleotides Enco
;  FILE REFERENCE: LEX-0298-USA
;  CURRENT APPLICATION NUMBER: US/10/044,807
;  CURRENT FILING DATE: 2002-01-11
;  PRIOR APPLICATION NUMBER: US 60/261,084
;  PRIOR FILING DATE: 2001-01-12
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: FastSEQ for Windows Version 4
; SEQ ID NO 1
;   LENGTH: 5289
;   TYPE: DNA
;   ORGANISM: homo sapiens
US-10-044-807-1

  Query Match             20.1%;  Score [illegible];  DB [illegible];  Length 5289;
  Best Local Similarity   60.8%;  Pred. No. [illegible];
  Matches 1742;  Conservative    0;  Mismatches [illegible];  Indels   45;  Gaps    [illegible];

Qy        186 [illegible]TTGGGG 245
Db         60 CCTGAGTTCAGGA[illegible] 119

Qy        246 CGACTGGAGTGACTGCTCCCTGACCTGTGGGGAGGAGCATCATA[illegible]GAGATG 305
Db        120 CCCATGGAGTGAATGCTCACGTACCTGCGGGGGTGGGGCCTCCTACTCTCTGAGGCGCTG 179

Qy        306 TTTGACTGGAAGGAATTGTGAAGGCAGAACATTCGGTACAAGACATGCAGCAATCATGA 365
Db        180 [illegible]GTGGA 239

Qy        366 CTGCCCTCCAGATGCAGAAGATTTCAGAGCCAGCAGTGCTCAGCCTACAATGATGTCCA 425
Db        240 [illegible]CATAATGATGTCAA 299

Qy        426 TTATCAGGGGATTACTATGAATGGCTTCCACGATATAATGATCTGCTGCCCCGTGTC 485
Db        300 GCACCATGGCCA[illegible] 359

Qy        486 ACTCAAGTGTCATGCACAAGGACAAAACTTGGTGGTGGAGCTGGCACTTAAGGTACTGGA 545
Db        360 ACTCAAGTGCCAAGCCAAAGGAACAACCCTGGTGGTGGAACTAGCACCTAAGGTCTAGA 419

Qy        546 TGGAACTGGTTGCAACACGGACTCCTTGGACATGGTGTATCAGTGGATCTGTCAGGCAGT 605
Db        420 TGGTACGCGTTGCTATACAGAA[illegible]AAATTGT 479

Qy        606 GGGCTGGGATCGGCAACTGGGAAGCAATGGCAAGGAGGACAACTGTGGAGTCTGTGCCGG 665
Db        480 TGGCTGCGATCACCAGCTGGGAAGCACCGTCAAGGAAGATAA[illegible]AACGG 539

Qy        666 CGATGGCTCCACCTGCAGGCTTGTACGGGGACAATCAAAGTCACAGCTTTCTCCTGAAAA 725
Db        540 AGATGGGTCCACCTGCCGGCTCGTCTGAGGGCAGTATAAA[illegible]AACCAA 599

Qy        726 AAGAGAAGAAAATGTAATTGCTGTTCCTTGGGAAGTGGAAGTGTGAGAATTACAGTGAA 785
Db        600 ATCGGATGATACTGTGGTTGCAATTCCCTATGGAAGTAGACATATTCGCCTTGTCTTAAA 659

Qy        786 AGGACCTGCCCACCTCTTTATTGAATCAAAAACACTTCAAGGAAGCAAAGGAGAACACAG 845
Db        660 AGG[illegible]GGGATAAAG[illegible]AAAACAG 719

Qy        846 CTTTAACAGCCCCGGCGTCTTTGTCGTAGAAAACACAACAGTGGAATTTCAGAGGGGCTC 905
Db        720 [illegible]AGAAATTCC 779

Qy        906 CGAGAGGCAAACTTTTAAGATTCCAGGACCTCTGATGGCTGATTTCATCTTCAAGACCAG 965
```

```
Db        780 AGACAAAGAGATACTGAGAATGGCTGGACCACTCACAGCAGATTTCATTGTCAAGATTCG 839

Qy        966 GTACACTGCAGCCAAAGACAGCGTGGTTCAGTTCTTCTTTTACCAGCCCATCAGTCATCA 1025

Db        840 TAACTCGGGCTCCGCTGACAGTACAGTCCAGTTCATCTTCTATCAACCCATCATCCACCG 899

Qy       1026 GTGGAGACAAACTGACTTCTTTCCCTGCACTGTGACGTGTGGAGGAGGTTATCAGCTCAA 1085

Db        900 ATGGAGGGAGACGGATTTCTTTCCTTGCTCAGCAACCTGTGGAGGAGGTTATCAGCTGAC 959

Qy       1086 TTCTGCTGAATGTGTGGATATCCGCTTGAAGAGGGTAGTTCCTGACCATTATTGTCACTA 1145

Db        960 ATCGGCTGAGTGCTACGATCTGAGGAGCAACCGTGTGGTTGCTGACCAATACTGTCACTA 1019

Qy       1146 CTACCCTGAAAATGTAAAACCAAAACCAAAACTGAACGAATGCAGCATGGATCCCTGCCC 1205

Db       1020 TTACCCAGAGAACATCAAACCCAAACCCAAGCTTCACGAGTGCAACTTGGATCCTTGTCC 1079

Qy       1206 ATCAAGTGATGGATTTAAAGAGATAATGCCCTATGACCACTTCCAACCTCTTCCTCGCTG 1265

Db       1080 AGCCAGTGACGGATACAAGCAGATCATGCCTTATGACCTCTACCATCCCCTTCCTCGGTG 1139

Qy       1266 GGAACATAATCCTTGGACTGCATGTTCCGTGTCCTGTGGAGGAGGGATTCAGAGACGGAG 1325

Db       1140 GGAGGCCACCCCATGGACCGCGTGCTCCTCCTCGTGTGGGGGGGCATCCAGAGCCGGGC 1199

Qy       1326 CTTTGTGTGTGTAGAGGAATCCATGCATGGAGAGATATTGCAGGTGGAAGAATGGAAGTG 1385

Db       1200 AGTTTCCTGTGTGGAGGAGGACATCCAGGGGCATGTCACTTCAGTGGAAGAGTGGAAATG 1259

Qy       1386 CATGTACGCACCCAAACCCAAGGTTATGCAAACTTGTAATCTGTTTGATTGCCCCAAGTG 1445

Db       1260 CATGTACACCCCTAAGATGCCCATCGCGCAGCCCTGCAACATTTTGACTGCCCTAAATG 1319

Qy       1446 GATTGCCATGGAGTGGTCTCAGTGCACAGTGACTTGTGGCCGAGGGTTACGGTACCGGGT 1505

Db       1320 GCTGGCACAGGAGTGGTCTCCGTGCACAGTGACATGTGGCCAGGGCCTCAGATACCGTGT 1379

Qy       1506 TGTTCTGTGTATTAACCACCGCGGAGAGCATGTTGGGGGCTGCAATCCACAACTGAAGTT 1565

Db       1380 GGTCCTCTGCATCGACCATCGAGGAATGCACACAGGAGGCTGTAGCCCAAAAACAAAGCC 1439

Qy       1566 ACACATCAAAGAAGAATGTGTCATTCCCATCCCGTGTTATAAACCAAAAGAAAAAGTCC 1625

Db       1440 CCACATAAAAGAGGAATGCATCGTACCCACTCCCTGCTATAAACCCAAAGAGAAACTTCC 1499

Qy       1626 AGTGGAAGCAAAATTGCCTTGGCTGAAACAAGCACAAGAACTAGAAGAGACCAGAATAGC 1685

Db       1500 AGTCGAGGCCAAGTTGCCATGGTTCAAACAAGCTCAAGAGCTAGAAGAAGGAGCTGCTGT 1559

Qy       1686 AACAGAAGAACCAACGTTCATTCCAGAACCCTGGTCAGCCTGCAGTACCACGTGTGGGCC 1745

Db       1560 GTCAGAGGAGCCCTCGTTCATCCCAGAGGCCTGGTCGGCCTGCACAGTCACCTGTGGTGT 1619

Qy       1746 AGGTGTGCAGGTCCGCGAGGTGAAGTGCCGTGTGCTCCTCACATTCACGCAGACTGAGAC 1805

Db       1620 GGGGACCCAGGTGCGAATAGTCAGGTGCCAGGTGCTCCTGTCTTTCTCTCAGTCCGTGGC 1679

Qy       1806 TGAGCTGCCCGAGGAAGAGTGTGAAGGCCCCAAGCTGCCCACC---------------GA 1850

Db       1680 TGACCTGCCTATTGACGAGTGTGAAGGGCCCAAGCCAGCATCCCAGCGTGCCTGTTATGC 1739

Qy       1851 ACGGCCCTGCCTCCTGGAAGCATGTGATGAGAGCCCGGCCTCCCGAGAGCTAGACATCCC 1910

Db       1740 AGGCCCATGCAGCGGGGAAATTCCTGAGTTCAACCCAGACGAGACAGATGGGCTCTTTGG 1799

Qy       1911 TCTCCCTGAGGACAGTGAGACGACTTACGACTGGGAGTACGCTGGGTTCACCCCTTGCAC 1970

Db       1800 TGGCCTGCAGGATTTCGACGAGCTGTATGACTGGGACTATGAGGGGTTCACCAAGTGCTC 1859

Qy       1971 AGCAACATGCTTGGGAGGCCATCAAGAAGCCATAGCAGTGTGCTTACATATCCAGACCCA 2030
```

```
Db       1860 CGAGTCCTGTGGAGGAGGTGTCCAGGAGGCTGTGGTGAGCTGCTTGAACAAACAGACTCG 1919

Qy       2031 GCAGACAGTCAATGACAGCTTGTGTGATATGGTCCACCGTCCTCCAGCCATGAGCCAGGC 2090

Db       1920 GGAGCCTGCTGAGGAGAACCTGTGCGTGACCAGCCGCCGGCCCCCACAGCTCCTGAAGTC 1979

Qy       2091 CTGTAACACAGAGCCCTGTCCCCCCAGGTGGCATGTGGGCTCTTGGGGGCCCTGCTCAGC 2150

Db       1980 CTGCAATTTGGATCCCTGCCCAGCAAGGTGGGAAATTGGCAAGTGGAGTCCATGTAGTCT 2039

Qy       2151 TACCTGTGGAGTTGGAATTCAGACCCGAGATGTGTACTGCCTGCACC------------- 2197

Db       2040 CACATGTGGGGTCGGCCTACAGACCAGAGACGTCTTCTGCAGCCACCTGCTTTCCAGAGA 2099

Qy       2198 -----CAGGGGACACCCCTGCCCCTCCTGAGGAGTCCCGACATGAAAACCCCCATGCTTT 2252

Db       2100 GATGAATGAAACAGTCATCCTGGCTGATGAGCTGTGTCGGCAGCCCAAGCCCAGCACGGT 2159

Qy       2253 ACAAGCATGCAATCAGTTTGACTGCCCTCCTGGCTGGCACATTGAAGAATGGCAGCAGTG 2312

Db       2160 GCAAGCTTGTAACCGCTTTAATTGCCCCCCAGCCTGGTACCCTGCACAGTGGCAGCCGTG 2219

Qy       2313 TTCCAGGACTTGTGGCGGGGGAACTCAGAACAGAAGAGTCACCTGTCGGCAGCTGCTAAC 2372

Db       2220 TTCCAGAACGTGTGGCGGGGGTGTTCAGAAACGTGAGGTTCTTTGCAAGCAGGCATGGC 2279

Qy       2373 GGATGGCAGCTTTTTGAATCTCTCAGATGAATTGTGCCAAGGACCCAAGGCATCGTCTCA 2432

Db       2280 TGATGGCAGCTTCCTGGAGCTTCCTGAGACCTTCTGTTCAGCTTCAAAACCTGCCTGCCA 2339

Qy       2433 CAAGTCCTGTGCCAGGACAGACTGTCCTCCACATTTAGCTGTGGGAGACTGGTCGAAGTG 2492

Db       2340 GCAAGCATGCAAGAAAGATGACTGTCCCAGCGAGTGGCTTCTCTCAGACTGGACAGAGTG 2399

Qy       2493 TTCTGTCACTTCTCGTGTTCGAATCCAGAGAACAAACCACCTCTGTCAAACCCTGCCACC 2552

Db       2400 TTCCACAAGCTGCGGGGAAGGCACCCAGACTCGAAGCGCCATTTGCCGAAAGATGCTGAA 2459

Qy       2553 CAAAGGTCGGCGCATCCCCCTCAGTCAGATCATCTGCACCGATCTACCAGGCTTCCCTCT 2612

Db       2460 AACCGGCCTCTCAACGCTTGTCAATTCCACCCTGTGCCCGCCCCTGCCTTTCTCTTCCTC 2519

Qy       2613 TGTAAGATCTTGCCAGATGCCTGAGTGCAGTAAAATCAAATCAGAGATGAAGACAAAACT 2672

Db       2520 CATCAGGCCCTGTATGCTGGC------------AACCTGTGCAAGGCCCGGGCGGCCATC 2567

Qy       2673 TGGTGAGCAGGGTCCGCAGATCCTCAGTGTCCAGAGAGTCTACATTCAGACAAGGCAAGA 2732

Db       2568 CACGAAGCACAGGCGGCACATCGGCGGCGCAGGAAGGTCTACATCCAGACTCGGCAGGCA 2627

Qy       2733 GAAGCGTATTAACCTGACCATTGGTAGCAGAGCCTATTTGCTGCCCAACACATCCGTGAT 2792

Db       2628 GAGGAAGCTGCACTTCGTGGTGGGGGGCTTCGCCTACCTGCTCCCCAAGACGGCGGTGGT 2687

Qy       2793 TATTAAGTGCCCCGTGCGACGATTCCAGAAATCTCTGATCCAGTGGGAGAAGGATGGCCG 2852

Db       2688 GCTGCGCTGCCCGGCGCGCAGGGTCCGCAAGCCCCTCATCACCTGGGAGAAGGACGGCCA 2747

Qy       2853 TTGCCTGCAGAACTCCAAACGGCTTGGCATCACCAAGTCAGGCTCACTAAAAATCCACGG 2912

Db       2748 GCACCTCATCAGCTCGACGCACGTCACGGTGGCCCCCTTCGGCTATCTCAAGATCCACCG 2807

Qy       2913 TCTTGCTGCCCCCGACATCGGCGTGTACCGGTGCATTGCAGGCTCTGCACAGGAAACAGT 2972

Db       2808 CCTCAAGCCCTCGGATGCAGGCGTCTACACCTGCTCAGCGGGCCCGGCCGGGAGCACTT 2867

Qy       2973 TGTGCTCAAGCTCATTGGTACTGACAACGGCTCATCGCACGGCCC 3017

Db       2868 TGTGATTAAGCTCATCGGAGGCAACCGCAAGCTCGTGGCCCGGCC 2912


RESULT 10
US-09-784-358-5
```

```
; Sequence 5, Application US/09784358
; Patent No. US20020006672A1
; GENERAL INFORMATION:
;  APPLICANT: Donoho, Gregory
;  APPLICANT:  Scoville, John
;  APPLICANT:  Turner, C. Alexander Jr.
;  APPLICANT:  Friedrich, Glenn
;  APPLICANT:  Zambrowicz, Brian
;  APPLICANT:  Sands, Arthur T.
;  TITLE OF INVENTION: NOVEL HUMAN THROMBOSPONDIN REPEAT PROTEINS AND
;  TITLE OF INVENTION:  POLYNUCLEOTIDES ENCODING THE SAME
;  FILE REFERENCE: LEX-0141-USA
;  CURRENT APPLICATION NUMBER: US/09/784,358
;  CURRENT FILING DATE:  2001-02-15
;  PRIOR APPLICATION NUMBER: US 60/183,282
;  PRIOR FILING DATE: 2000-02-17
;  NUMBER OF SEQ ID NOS: 17
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
;   LENGTH: 1119
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-784-358-5

  Query Match             20.0%;  Score 1017;  DB 10;  Length 1119;
  Best Local Similarity   99.8%;  Pred. No. 1.5e-294;
  Matches 1029;  Conservative    0;  Mismatches    0;  Indels    2;  Gaps    1;

Qy        311 CTGGAAGGAATTGTGAAGGGCAGAACATTCGGTACAAGACATGCAGCAATCATGACTGCC 370
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         89 CTGGAAGGAATTGTGAAGGGCAGAACATTCGGTACAAGACATGCAGCAATCATGACTGCC 148

Qy        371 CTCCAGATGCAGAAGATTTCAGAGCCCAGCAGTGCTCAGCCTACAATGATGTCCAGTATC 430
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        149 CTCCAGATGCAGAAGATTTCAGAGCCCAGCAGTGCTCAGCCTACAATGATGTCCAGTATC 208

Qy        431 ACGGGCATTACTATGAATGGTTTCCACGATATAATGATCCTGCTGCCCCGTGTGCACTCA 490
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        209 ACGGGCATTACTATGAATGGTTTCCACGATATAATGATCCTGCTGCCCCGTGTGCACTCA 268

Qy        491 AGTGTCATGCACAAGGACAAAAACTTGGTGGTGGAGCTGGCACCTAAGGTACTGGATGAA 550
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        269 AGTGTCATGCACAAGGACAAAAACTTGGTGGTGGAGCTGGCACCTAAGGTACTGGATGAA 328

Qy        551 CTGGTTGCAACACGGACTTCTGGACATGTGTATCAGTGGCATCTGGTCAGGCAGTGGCCT 610
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        329 CTGGTTGCAACACGGACTTCTGGACATGTGTATCAGTGGCATCTGGTCAGGCAGTGGCCT 388

Qy        611 GGCATGGGCAACTGGGAAGCAATGCCAAGGAGGACAACTGTGGAGTCTGTGCCGGGGATG 670
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        389 GGCATGGGCAACTGGGAAGCAATGCCAAGGAGGACAACTGTGGAGTCTGTGCCGGGGATG 448

Qy        671 GCTCCACCTGCAGGCTTGTAGGGGGACAATCAAAGTCACACGTTTCTCCTGAAAAAAGAG 730
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        449 GCTCCACCTGCAGGCTTGTAGGGGGACAATCAAAGTCACACGTTTCTCCTGAAAAAAGAG 508

Qy        731 AAGAAAATGTAATTGGTTGTTCTTTGGGAAGTGAAGTGTGAGAATTACAGTGAAAGGAC 790
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        509 AAGAAAATGTAATTGGTTGTTCTTTGGGAAGTGAAGTGTGAGAATTACAGTGAAAGGAC 568

Qy        791 CTGCCCACCTCTTTATTGAATCAAAAACACTTAAGGAAGCAAAGGAGAACACAGTTTTA 850
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        569 CTGCCCACCTCTTTATTGAATCAAAAACACTTAAGGAAGCAAAGGAGAACACAGTTTTA 628

Qy        851 ACACGCCCCGGGTGTTTTGGTTAGAAAACACAAAGTGGAATTTCAGAGGGGTCCGAGA 910
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        629 ACACGCCCCGGGTGTTTTGGTTAGAAAACACAAAGTGGAATTTCAGAGGGGTCCGAGA 688

Qy        911 GGCAAACTTTTAAGATTCCAGGACCTCTGATGGTGATTTCATCTTCAAGACCAGGTACA 970
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        689 GGCAAACTTTTAAGATTCCAGGACCTCTGATGGTGATTTCATCTTCAAGACCAGGTACA 748

Qy        971 CTGCAGCCAAAGACACCTTGGTTCATTTCTTTTACCACCCATCAGTCATCATTGGA 1030
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        749 CTGCAGCCAAAGACACCTTGGTTCATTTCTTTTACCACCCATCAGTCATCATTGGA 808

Qy       1031 GACAAACTGACTTTTTCCTTGCACTTCACGTGTGGAGGAGGTTATCAGCTCAATTCTG 1090
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        809 GACAAACTGACTTTTTCCTTGCACTTCACGTGTGGAGGAGGTTATCAGCTCAATTCTG 868

Qy       1091 CTGAATGGTGTCATATCCGCTTGAAGAGGGTAGTTCCTGACCATTATTGTCACTACTACC 1150
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        869 CTGAATGGTGTCATATCCGCTTGAAGAGGGTAGTTCCTGACCATTATTGTCACTACTACC 928

Qy       1151 CTGAAAATGTAAAACCAAAACCAAAACTGAAGGAATGCAGCATGCATCCTGCCCATCAA 1210
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        929 CTGAAAATGTAAAACCAAAACCAAAACTGAAGGAATGCAGCATGCATCCTGCCCATCAA 988

Qy       1211 GTGATGGATTTAAAGAGAAAATGCCCTATGACCACTTCCAACCTCTTCCTC--GCTGGGA 1268
              |||||||||||||||||||||||||||||||||||||||||||||||||||  |||||||
Db        989 GTGATGGATTTAAAGAGAAAATGCCCTATGACCACTTCCAACCTCTTCCTCCGAGCTGGGA 1048

Qy       1269 ACATAATCCTTGGACTGCATGTTCCTGCCCTTCTGGAGGAGGGATTCAGAGACGGAGCTT 1328
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1049 ACATAATCCTTGGACTGCATGTTCCTGCCCTTCTGGAGGAGGGATTCAGAGACGGAGCTT 1108

Qy       1329 TGTGTGTGTAG 1339
              |||||||||||
Db       1109 TGTGTGTGTAG 1119


RESULT 11
US-09-764-853-134
; Sequence 134, Application US/09764853
; Patent No. US20020006672A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PZ706
;  CURRENT APPLICATION NUMBER: US/09/764,853
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - consult PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 939
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 134
;   LENGTH: 5769
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-764-853-134

  Query Match             18.2%;  Score 925.4;  DB 10;  Length 5769;
  Best Local Similarity   59.4%;  Pred. No. 1.7e-266;
  Matches 1701;  Conservative    1;  Mismatches 1067;  Indels   96;  Gaps    4;

Qy        186 CCAGACCTCAAGAAACACTCGTTCAGATGAAGACAAAGATGGCAACTGGGATGCTTGGGG 245
              || ||  || || | | | || || || || |||   || |||   |||||||| |||||
Db        124 CCTGAGTTCAGGACCGCAGCAGGAGGAGGAGCCGGACGGCCTATGGCATGCCTGGGG 183

Qy        246 CGACTGGAGTGACTGCTCCCGGACCTGTGGGGGAGGAGCATCATATTCTCTCGGAGATC 305
              |   ||||||| |||||  | |||| |||| ||    | |  ||||||  ||  | ||
Db        184 CCCATGGAGTGAATGCTCACGCACCTGCGGGGGTGGGCCTCCTACTCTCTAGGCGCTG 243

Qy        306 TTTGACTGGAAGGAATTGTGAAGGGCAGAACATTCGGTACAACACATGCAGCAATCATGA 365
                |||    |  |  ||   ||||||||      || || || ||||  |||||||| |||      ||
Db        244 CCTGAGCAGCAAGAGCTGTGAAGGAAGAAATACCCGATACAGAACATGCAGTAATGTGGA 303

Qy        366 CTGCCCTCCAGATGCAGAAGATTTCAGAGCCCAGCAGTGCTCAGCCTACAATGATGTCCA 425
              ||||| |||    ||||   |||||  ||||  ||||  |||||||    |  ||||||||| |
Db        304 CTGCCCACCAGAAGCAGGTGATTTCCGAGCTCAGCAATGCAGCTCATAATGATGTCAA 363

Qy        426 GTATCAGGGGCATTACTATGAATGGTTTCCACGATATAATGATCCTGCTGCCCCGTGTGC 485
              | |  || ||  || |   |||||||||||||      | ||||||  ||||      |||  ||||  |
Db        364 GCACCATGGCCAGTTTTATGAATGGCTTCCTGTGTTAAATGACCCTGACAACCCATGTTC 423

Qy        486 ACTCAAGTGTCATGCACAAGGACAAAAACTTGGTGGTGGAGCTGGCACCTAAGGTACTGGA 545
              |||||||||| |  ||  |||||  || |  |||  ||  | || ||  ||||||||||  | ||
Db      424 ACTCAAGTGCCAAGCCAAAGGAACAACCCTGGTTGTTGAACTAGCACCTAAGGTCTTAGA 483

Qy      546 TGGAACTCGTTGCAACACGGACTCCTTGGACATGTGTATCAGTGGCATCTGTCAGGCAGT 605
             ||| || |||||  |  || || ||||| ||||| ||||||||   | || ||     ||
Db      484 TGGTACGCGTTGCTATACAGAATCTTTGGATATGTGCATCAGTGGTTTATGCCAAATTGT 543

Qy      606 GGGCTGCGATCGGCAACTGGGAAGCAATGCCAAGGAGGACAACTGTGGAGTCTGTGCCGG 665
              ||||||||||   || ||||||||||   | |||||| || |||||||| |||||   |||
Db      544 TGGCTGCGATCACCAGCTGGGAAGCACCGTCAAGGAAGATAACTGTGGGGTCTGCAACGG 603

Qy      666 CGATGGCTCCACCTGCAGGCTTGTACGGGGACAATCAAAGTCACACGTTTCTCCTGAAAA 725
              ||||| ||||||||| |||| || || || || |   || || ||   | ||  |     ||
Db      604 AGATGGGTCCACCTGCCGGCTGGTCCGAGGGCAGTATAAATCCCAGCTCTCCGCAACCAA 663

Qy      726 AAGAGAAGAAAATGTAATTGCTGTTCCTTTGGGAAGTCCAAGTGTGAGAATTACAGTGAA 785
             |    || || | | |    | ||   ||| |   |||||| ||  |  |  |  ||     | ||
Db      664 ATCGGATGATACTGTGGTTGCAATTCCCTATGGAAGTAGACATATTCGCCTTGTCTTAAA 723

Qy      786 AGGACCTGCCCACTTCTTTATTGAATCAAAAACACTTCAAGGAAGCAAAGGAGAACACAG 845
             ||| |||  ||| | | |  ||| | ||||| || || || |  ||||| ||| ||||
Db      724 AGGTCCTGATCACTTATATCTGGAAACCAAAACCCTCCAGGGGACTAAAGGTGAAAACAG 783

Qy      846 CTTTAACAGCCCCGGCGTCTTTGTCGTAGAAAACACAACAGTGGAATTTCAGAGGGGCTC 905
               | | |  | | ||    ||  | || || ||  | |  ||||| || ||||        |
Db      784 TCTCAGCTCCACAGGAACTTTCCTTGTGGACAATTCTAGTGTGGACTTCCAGAAATTTCC 843

Qy      906 CGAGAGGCAAACTTTTAAGATTCCAGGACCTCTGATGGCTGATTTCATCTTCAAGACCAG 965
             || |   | |    | |  ||  | |||||| || |   || ||||||||  ||||||   |
Db      844 AGACAAAGAGATACTGAGAATGGCTGGACCACTCACAGCAGATTTCATTGTCAAGATTCG 903

Qy      966 GTACACTGCAGCCAAAGACAGCGTGGTTCAGTTCTTCTTTACCAGCCCATCAGTCATCA 1025
              || | |   ||   |||||      || |||||| |||| || || |||||||  ||| |
Db      904 TAACTCGGGCTCCGCTGACAGTACAGTCCAGTTCATCTTCTATCAACCCATCATCCACCG 963

Qy     1026 GTGGAGACAAACTGACTTCTTTCCCTGCACTGTGACGTGTGGAGGAGGTTATCAGCTCAA 1085
             |||||   | || || ||||||||| ||| | |   || |||||||||||||||||||| |
Db      964 ATGGAGGGAGACGGATTTCTTTCCTTGCTCAGCAACCTGTGGAGGAGGTTATCAGCTGAC 1023

Qy     1086 TTCTGCTGAATGTGTGGATATCCGCTTGAAGAGGGTAGTTCCTGACCATTATTGTCACTA 1145
             || ||||| ||     ||| | |     ||  | || ||| |||||||| || ||||||||
Db     1024 ATCGGCTGAGTGCTACGATCTGAGGAGCAACCGTGTGGTTGCTGACCAATACTGTCACTA 1083

Qy     1146 CTACCCTGAAAATGTAAAACCAAAACCAAAACTGAAGGAATGCAGCATGGATCCCTGCCC 1205
             ||||| || || |   | ||||| ||||| || ||  |||| |||| | ||||||| || ||
Db     1084 TTACCCAGAGAACATCAAACCCAAACCCAAGCTTCAGGAGTGCAACTTGGATCCTTGTCC 1143

Qy     1206 ATCAAGTGATGGATTTAAAGAGATAATGCCCTATGACCACTTCCAACCTCTTCCTCGCTG 1265
            |  | |                                                 | ||
Db     1144 AGCCA---------------------------------------------------GGTG 1152

Qy     1266 GGAACATAATCCTTGGACTGCATGTTCCGTGTCCTGTGGAGGAGGCATTCAGAGACGGAG 1325
            |||     |  ||  |||||| || || |||    || ||||| || || || ||||| |||
Db     1153 GGAGGCCACCCCATGGACCGCGTGCTCCTCCTCGTGTGGGGGGGCATCCAGAGCCGGGC 1212

Qy     1326 CTTTGTGTGTGTAGAGGAATCCATGCATGGAGAGATATTGCAGGTGGAAGAATGGAAGTG 1385
              ||    ||||| |||||    ||| || ||   |   |         |||||||| ||||| ||
Db     1213 AGTTTCCTGTGTGGAGGAGGACATCCAGGGGCATGTCACTTCAGTGGAAGAGTGGAAATG 1272

Qy     1386 CATGTACGCACCCAAACCCAAGGTTATGCAAACTTGTAATCTGTTTGATTGCCCCAAGTG 1445
            ||||||| ||  ||          |    |||   | || ||   | ||||| ||||| || ||
Db     1273 CATGTACACCCCTAAGATGCCCATCGCGCAGCCCTGCAACATTTTTGACTGCCCTAAATG 1332

Qy     1446 GATTGCCATGGAGTGGTCTCAGTGCACAGTGACTTGTGGCCGAGGGTTACGGTACCGGGT 1505
            |  |  ||    |||||||||| ||||||||||| ||||||| ||   |   | ||||| ||
Db     1333 GCTGGCACAGGAGTGGTCTCCGTGCACAGTGACATGTGGCCAGGGCCTCAGATACCGTGT 1392

Qy     1506 TGTTCTGTGTATTAACCACCGCGGAGAGCATGTTGGGGGCTGCAATCCACAACTGAAGTT 1565
              || || || ||  |||| || |||   |||      || |||||  |    ||| ||    |||
Db     1393 GGTCCTCTGCATCGACCATCGAGGAATGCACACAGGAGGCTGTAGCCCAAAAACAAAGCC 1452

Qy     1566 ACACATCAAAGAAGAATGTGTCATTCCCATCCCGTCTTATAAACCAAAAGAAAAAAGTCC 1625
             ||||| ||||| |||  | | |||  || || |||||||| ||||| |||   |||
Db     1453 CCACATAAAAGAGAATGCATCGTACCACTGCTGCTATAAACCCAAAGACAAACTTCC 1512

Qy     1626 AGTGGAAGCAAAATTGCCTTGGCTGAAACAAGCACAAGAACTAGAAGAGACCAGAATAGC 1685
            ||| || || || ||||| ||| | ||||||| ||||| ||||||||             |
Db     1513 AGTCGAGGCCAAGTTGCCATGGTTCAAACAAGCTCAAGAGCTAGAAGAAGGAGCTGCTGT 1572

Qy     1686 AACAGAAGAACCAACGTTCATTCCAGAACCCTGGTCAGCCTGCAGTACCACGTGTGGCC 1745
              |||| || ||  ||||||| |||:|  ||||||| |||||||     ||| |||||
Db     1573 GTCAGAGGAGCCCTCGTTCATCCCARAGGCCTGGTCGGCCTGCACAGTCACCTGTGGTGT 1632

Qy     1746 AGGTGTGCAGGTCCGCGAGGTGAAGTGCCGTGTGCTCCTCACATTCACGCAGACTGAGAC 1805
             ||     ||||| ||     || | |||||  |||||||   | ||| | ||| | | | |
Db     1633 GGGGACCCAGGTGCGAATAGTCAGGTGCCAGGTGCTCCTGTCTTTCTCTCAGTCCGTGGC 1692

Qy     1806 TGAGCTGCCCGAGGAAGAGTGTGAAGGCCCCAAGCTGCCCACC---------------GA 1850
            ||| |||||     || |||||||||||| ||||||     |  ||                |
Db     1693 TGACCTGCCTATTGACGAGTGTGAAGGGCCCAAGCCAGCATCCCAGCGTGCCTGTTATGC 1752

Qy     1851 ACGGCCCTGCCTCCTGGAAGCATGTGATGAGAGCCCGGCCTCCCGAGAGCTAGACATCCC 1910
            | | || |||  |   ||||      |||      | ||| | |       |||        | |
Db     1753 AGGCCCATGCAGCGGGGAAATTCCTGAGTTCAACCCAGACGAGACAGATGGGCTCTTTGG 1812

Qy     1911 TCTCCCTGAGGACAGTGAGACGACTTACGACTGGGAGTACGCTGGGTTCACCCCTTGCAC 1970
            |  ||   ||||    ||    |    || ||||||||||| |   ||||||||||     ||| |
Db     1813 TGGCCTGCAGGATTTCGACGAGCTGTATGACTGGGAGTATGAGGGGTTCACCAAGTGCTC 1872

Qy     1971 AGCAACATGCTTGGGAGGCCATCAAGAAGCCATAGCAGTGTGCTTACATATCCAGACCCA 2030
             |   | ||     |||||     || || ||   | |       ||||| | |  |||||| |
Db     1873 CGAGTCCTGTGGAGGAGGTGTCCAGGAGGCTGTGGTGAGCTGCTTGAACAAACAGACTCG 1932

Qy     2031 GCAGACAGTCAATGACAGCTTGTGTGATATGGTCCACCGTCCTCCAGCCATGAGCCAGGC 2090
            | || | |    | || | | |||| |  |      || ||| || |||      |      || |
Db     1933 GGAGCCTGCTGAGGAGAACCTGTGCGTGACCAGCCGCCGGCCCCCACAGCTCCTGAAGTC 1992

Qy     2091 CTGTAACACAGAGCCCTGTCCCCCCAGGTGGCATGTGGGCTCTTGGGGGCCCTGCTCAGC 2150
            ||| ||     || ||||| ||  | |||||| |   | |||     ||| | || ||
Db     1993 CTGCAATTTGGATCCCTGCCCAGCAAGGTGGGAAATTGGCAAGTGGAGTCCATGTAGTCT 2052

Qy     2151 TACCTGTGGAGTTGGAATTCAGACCCGAGATGTGTACTGCCTGCACC------------- 2197
              || ||||| || ||   | |||||| |||| || | ||||    ||||
Db     2053 CACATGTGGGGTCGGCCTACAGACCAGAGACGTCTTCTGCAGCCACCTGCTTTCCAGAGA 2112

Qy     2198 -----CAGGGGAGACCCCTGCCCCTCCTGAGGAGTGCCGAGATGAAAAGCCCATGCTTT 2252
                    |       |        ||   ||||  ||||  ||    |      ||||||      |   |
Db     2113 GATGAATGAAACAGTCATCCTGGCTGATGAGCTGTGTCGCCAGCCCAAGCCCAGCACGGT 2172

Qy     2253 ACAAGCATGCAATCAGTTTGACTGCCCTCCCTGGCTGGCACATTGAAGAATGGCAGCAGTG 2312
              ||||| || || |   ||| | ||||| || | |||| ||   || | | ||||||| |||
Db     2173 GCAAGCTTGTAACCGGCTTTAATTGCCCCCCAGCCTGGTACCCTGCACAGTGGCAGCCGTG 2232

Qy     2313 TTCCAGGACTTGTGGCGGGGGAACTCAGAACAGAAGAGTCACCTGTCGGCAGCTGCTAAC 2372
            |||||| || ||||||||||      |||||| |     ||     ||    |||||     | |
Db     2233 TTCCAGAACGTGTGGCGGGGGTGTTCAGAAACGTGAGGTTCTTTGCAAGCAGCGCATGGC 2292

Qy     2373 GCATGGCAGCTTTTTCAATCTCTCAGATGAATTGTGCCAAGGACCCAAGGCATCGTCTCA 2432
             |||||||||||  || | ||   | ||     || ||    ||     | ||  |  |  | ||
Db     2293 TCATGGCAGCTTCCTGGAGCTTCCTGACACCTTCTGTTCAGCTTCAAAACCTGCCTGCCA 2352

Qy     2433 CAAGTCCTGTGCCAGGACAGACTGTCCTCCACATTTAGCTGTGGGAGACTGGTCGAAGTG 2492
              |   | ||    |       ||||||||      | |     | |    ||||||| |  ||||
Db     2353 GCAAGCATGCAAGAAAGATGACTGTCCCAGCGAGTGGCTTCTCTTCAGACTGGACAGAGTG 2412

Qy     2493 TTCTGTCAGTTGTGGTGTTGGAATCCAGAGAAGAAAGCAGGTGTGTCAAAGGCTGGCAGC 2552
            |||       || || || |   || | |||||     |||          | || | || | ||
Db     2413 TTCCACAAGCTGCGGGGAAGGCACCCAGACTCGAAGCGCCATTTGCCGAAAGATGCTGAA 2472

Qy     2553 CAAAGGTCGGCGCATCCCCCTCAGTGAGATGATGTGCAGGGATCTACCAGGGTTCCCTCT 2612
              | || |     |         ||| |   |    |||||  |    || ||     |    |
Db     2473 AACCGGCCTCTCAACGGTTGTCAATTCCACCCTGTGCCCGCCCCTGCCTTTCTCTTCCTC 2532

Qy     2613 TGTAACATCTTGCCAGATGCCTGAGTGCAGTAAAATCAAATCAGAGATGAAGACAAAACT 2672
              | || |  ||    | ||  |                  || |    || |      |     |
Db     2533 CATCAGGCCCTGTATGCTGGC            AACCTGTGCAAGGCCCGGGCGGCCATC 2580
```

```
Qy       2673 TGTTGAGCAGGTTCCGCAGATCCTCAGTGTCCAGAGAGTCTACATTCAGACAAGCGAAGA 2732
Db       2581 CAGGAAGCACAGCCCGCACATCCGCGCCCGCAGGAAGGTCTACATACAGACCTCCAGGCA 2640

Qy       2733 GAAGCGTATTAACCTGACCATTGGTAGCAGAGCCTATTTGCTGCCCAACACATCCGTGAT 2792
Db       2641 GAGGAAGCTGCACTTCGTCCCCGGGGGCTTCGCCTACCTGCTCCCCAAGACGGCCGCGGT 2700

Qy       2793 TATTAAGTGCCCCGTGGCGACGATTCCAGAAATCCCTGATCCAGTGGGAGAAGGATGGCCG 2852
Db       2701 GCTGCGCTGCCCGGCGGCGCAGGCTCCGCAAGCCCCTCATCACCTGGGAGAAGGACGGCCA 2760

Qy       2853 TTGCCTGCAGAACTCCAAACGGCTTGGCATCACCAAGTCAGGCTCACTAAAAATCCACGG 2912
Db       2761 GCACCTCATCACCTTGACGCACGTCACGGTGGCCCCCTTCGGCTATCTCAAGATCCACCG 2820

Qy       2913 TCTTGCTGCCCCGACATCGCGGTGTACCGGTGCATTGCAGGCTCTGCACAGGAAACAGT 2972
Db       2821 CCTCAAGCCCTCGGATGCAGGCGTCTACACCTGCTCAGCGGGCCCGGCCCGGAGCACTT 2880

Qy       2973 TGTGCTCAAGTTCATCGGCTACCGACAACCGGCTCATCGCACGCCC 3017
Db       2881 TGTGATTAAGCTCATCGGAGTTAACCGAAGCCCCGGGCCCGCC 2925


RESULT 12
US-09-800-729-18
; Sequence 18, Application US/09800729
; Patent No. US20020068310A1
; GENERAL INFORMATION:
;  APPLICANT: Ni et al.
;  TITLE OF INVENTION: 32 Human secreted proteins
;  FILE REFERENCE: PZ044P1
;  CURRENT APPLICATION NUMBER: US/09/800,729
;  CURRENT FILING DATE: 2001-03-08
;  PRIOR APPLICATION NUMBER: PCT/US00/26013
;  PRIOR FILING DATE: 2000-09-22
;  PRIOR APPLICATION NUMBER: 60/155,709
;  PRIOR FILING DATE: 1999-09-24
;  NUMBER OF SEQ ID NOS: 217
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
;   LENGTH: 5720
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-800-729-18

  Query Match             18.2%;  Score 924.2;  DB 10;  Length 5720;
  Best Local Similarity   59.4%;  Pred. No. 3.9e-266;
  Matches 1701;  Conservative    0;  Mismatches 1068;  Indels   96;  Gaps    4;

Qy        186 CCAGACCTCAAGAAACACTCCTTCAGATGAAGACAAAGATGCCAACTGGGATGCTTCCGG 245
Db        126 CCTGAGTTCCAGGACCGCACCCCCGAGGAGGACGGGACGGCTATGGATGCCTGGGG 185

Qy        246 CGACTGGAGTGACTGCTCCCGGACCTGTGGGGGAGGAGCATCATATTCTCTGCGGAGATG 305
Db        186 CCCATGGAGGGAACCCCACCACCTGGGGGGGGGCTCCTACTCTCTGAGGCGCTG 245

Qy        306 TTTGATTGGAAGGAATTGTGAAGGGTAGAACATTCCGTACAAGACATGCAGCAATTATTGA 365
Db        246 CCTGAGCAGCAAGAGCTGTGAAGGAAGAAATATCCGATACAGAACATGCAGTAATGTGGA 305

Qy        366 CTGCCCTCCAGATGCAGAAGATTTCAGAGCCCAGCAGTGCTCAGCCTACAATGATGTCCA 425
Db        306 CTGCCCACCAGAAGCAGGTGATTTCCGAGCTCAGCAATGCTCAGCTCATAATGATGTCAA 365

Qy        426 GTATCAGGGGCATTACTATGAATGGCTTCCACGATATAATGATCCTGCTGCCCCGTGCGC 485
Db        366 GCACCATGGCCAGTTTTATGAATGGCTTCCTGTGTCTAATGACCCTGACAACCCATGCTC 425

Qy        486 ACTCAAGTGTCATGCACAAGGACAAAACTTGGTGGTGGAGCTGGCACCTAAGGTACTCGA 545
```

```
Db        426 ACTCAAGTGCAAGCCAAAGGAACAACCTGGTTGGTGAACTAGCACCTAAGGTCTTAGA 485

Qy        546 TGGAACTCGTTGCAACACGGACTCCTTGGACATGTGTATCAGTGGCATCTGCCAGGCAGT 605
Db        486 TGGTACGCGTTGCTATACAGAATCTTTGGATATGTGCATCAGTGGTTTATGCCAAATTGT 545

Qy        606 GGGCTGCGATCGGCAACTGGGAAGCAATGCCAAGGAGGACAACTGTGGAGTCTGTGCCGG 665
Db        546 TGGCTGCGATCACCAGCTGGGAAGCACCGTCAAGGAAGATAACTGTGGGGTCTGCAACGG 605

Qy        666 CGATGGCTCCACCTGCAGGCTTGTACGGGGACAATCAAAGTCACACGTTTCTCCTGAAAA 725
Db        606 AGATGGGTCCACCTGCCGGCTGGTCCGAGGGCAGTATAAATCCCAGCTCTCCGCAACCAA 665

Qy        726 AAGAGAAGAAAATGTAATTGCTGTTCCTTGGGAAGTCGAAGCGTGAGAATCACAGTGAA 785
Db        666 ATCGGATGATACTGTGGTTGCAATTCCCTATGGAAGTAGACATATTCGCCTTGTCTTAAA 725

Qy        786 AGGACCTGCCTACCTCTTTATTGAATCAAAAACACTTCAAGGAAGCAAAGGAGAACACAG 845
Db        726 AGGTCCTGATCACTTATATCTGGAAACCAAAACCCTCCAGGGGACTAAAGGGGAAAACAG 785

Qy        846 CTTTAACAGCCCTGGCGTCTTTGTCGTAGAAAACACAACAGTGAATTCCAGAGGGGCTC 905
Db        786 TCTCAGCTCCACAGGAACTTTCTTGTGGACAATTCTAGTGGGACTTCCAGAAATTTCC 845

Qy        906 CGAGAGGCAAACTTTTAAGATTCCAGGACCTCTGATGGCTGATTTCATCTTCAAGACCAG 965
Db        846 AGACAAAGAGATACTGAGAATGGCTGGACCACTCACAGCAGATTTCATCGTCAAGATTCG 905

Qy        966 GTACACTGCAGCTAAAGACAGCGTGGTTCAGTTCTTCTTTACCAGCCCATCAGTCATCA 1025
Db        906 TAACTCGGGCTCCGCTGACAGTACAGTCCAGTTCATCTTCTATCAACCCATCATCCACCG 965

Qy       1026 GTGGAGACAAACTGACTTCTTTCCCTGCACTGTGACGTGTGGAGGAGGTTATCAGCTCAA 1085
Db        966 ATGGAGGGAGACGGATTTCTTTCCTTGCTCAGCAACCTGTGGAGGAGGTTATCAGCTGAC 1025

Qy       1086 TTCTGCTGAATGTGTGGATATCCGCTTGAAGAGGGTAGCTCCTGACCATTATCTGTCACTA 1145
Db       1026 ATCGGCTGAGTGCTACGATCTGAGGAGCAACCGTGTGGTTGCTGACCAATACTGTCACTA 1085

Qy       1146 CTACCCTGAAAATGTAAAACCAAAACCAAAACTGAAGGAATGCAGCATGGATCCCTGCCC 1205
Db       1086 TTACCCAGAGAACATCAAACCCAAACCCAAGCTTCAGGAGTGCAACTTGGATCCTTGTCC 1145

Qy       1206 ATCAAGTGATGGATTTAAAGAGATAATGCCCTATGACCACTTCCAACCTCTTCCTCGGCTG 1265
Db       1146 AGCCA                                                    GGTG 1154

Qy       1266 GGAACATAATCCTTGGACTGCATGTTCCGTGTCCTGTGGAGGAGGGATTCAGAGACGGAG 1325
Db       1155 GGAGGCCACCCATGGACCGCGTGCTCCTCCTCGTGTGGGGGGGGATCCAGAGCCGGGC 1214

Qy       1326 CTTTGTGTGTGTAGAGGAATCCATGCATGGAGAGATATTGCAGGTGGAAGAATGGAAGTG 1385
Db       1215 AGTTTCCTGTGTGGAGGAGGACATCCAGGGGCATGTCACTTCAGTGGAAGAGTGGAAATG 1274

Qy       1386 CATGTACGCACCCAAACCCAAGGTTATGCAAACTTGTAATCTGTTTGATTGCCCCAAGTG 1445
Db       1275 CATGTACACCCCTAAGATGCCCATCGCGCAGCCTTGCAACATTTTTGACTGCCCTAAATG 1334

Qy       1446 GATTGCCATGGAGTGGTCTCAGTGCACAGTGACTTGTGGCCGAGGGTTACGGTACCGGGT 1505
Db       1335 GCTGGCACAGGAGTGGTCTCCGTGCACAGTGACATGTGGCCAGGGCCTCAGATACCGTGT 1394

Qy       1506 TGTTCTGTGTATTAACCACCGCGGAGAGCATGTTGGGGGCTGCAATCCACAACTGAAGTT 1565
Db       1395 GGTCCTCTGCATCGACCATCGAGGAATGCACACAGGAGGCTGTAGCCCAAAAACAAAGCC 1454

Qy       1566 ACACATCAAAGAAGAATGCGTCATCCCCATCCCGTGTTATAAACCAAAAGAAAAAAGTCC 1625
Db       1455 CCACATAAAAGAGGAATGCATCGTACCCACTCCCTGCTATAAACCCAAAGAGAAACTTCC 1514
```

```
Qy       1626 ACTCGAAGCAAAATTGCCTTGCCTGAAACAAGCACAAGAACTAGAAGAGACCAGAATAGC 1685
              || | || || || ||||| ||| | |||||||| ||||| ||||||||          |
Db       1515 AGTCGAGGCCAAGTTGCCATGGTTCAAACAAGCTCAAGAGCTAGAAGAAGGAGCTGCTGT 1574

Qy       1686 AACAGAAGAACCAACTTTCATTCCAGAACCCTGGTCAGCCTGCAGTACCACGTGTGGGCC 1745
                ||| || ||   || |||| ||| |   |||||| |||||||     ||| |||||
Db       1575 GTCAGACGAGCCCTCGTTCATCCCAAAGGCCTGGTCGGCCTGCACAGTCACCTGTGGTGT 1634

Qy       1746 AGGTGTGCAGGTCCGCGAGGTGAAGTGCCGTGTGCTCCTCACATTCACGCAGACTGAGAC 1805
               ||    ||||| ||     || | |||||   |||||||| | ||| | ||| | | | |
Db       1635 GGGGACCCAGGTGCGAATAGTCAGGTGCCAGGTGCTCCTGTCTTTCTCTCAGTCCGTGGC 1694

Qy       1806 TGAGCTGCCCGAGGAAGAGTGTGAAGGCCCCAAGCTGCCCACC---------------GA 1850
              ||    ||||    || ||||||||||| |||||||     |  ||             |
Db       1695 TGACCTGCCTATTGACGAGTGTGAAGGGCCCAAGCCAGCATCCCAGCGTGCCTGTTATGC 1754

Qy       1851 ACGGCCCTGCCTCCTGGAAGCATGTGATGAGAGCCCGGCCTCCCGAGAGCTAGACATCCC 1910
              | | || |||  |  ||||      |||    | ||| | |     |||       | |
Db       1755 AGGCCCATGCAGCGGGGAAATTCCTGAGTTCAACCCAGACGAGACAGATGGGCTCTTTGG 1814

Qy       1911 TCTCCCTGAGGACAGTGAGACGACTTACGACTGGGAGTACGCTGGGTTCACCCCTTGCAC 1970
              | ||   |||  |  |    || ||||||||||| |  ||||||||||    ||| |
Db       1815 TGGCCTGCAGGATTTCGACGAGCTGTATGACTGGGAGTATGAGGGGTTCACCAAGTGCTC 1874

Qy       1971 AGCAACATGCTTGGGAGGCCATCAAGAAGCCATAGCAGTGTGCTTACATATCCAGACCCA 2030
                |   | ||     |||||     || || ||   | |        ||||| | | ||||| |
Db       1875 CGAGTCCTGTGGAGGAGGTGTCCAGGAGGCTGTGGTGAGCTGCTTGAACAAACAGACTCG 1934

Qy       2031 GCAGACAGTCAATGACAGCTTGTGTGATATGGTCCACCGTCCTCCAGCCATGAGCCAGGC 2090
              | || | |    | || | | |||| |  |    || ||| || |||    |     || |
Db       1935 GGAGCCTGCTGAGGAGAACCTGTGCGTGACCAGCCGCCGGCCCCCACAGCTCCTGAAGTC 1994

Qy       2091 CTGTAACACAGAGCCCTGTCCCCCAGGTGGCATG GGGCTCTTGCCGCCCCTGCTCAGC 2150
              ||| ||    || ||||| ||  | |||||| |    |||   ||| | || ||
Db       1995 CTGCAATTTGGATCCCTGCCCAGCAAGGTGGGAAATTGGCAAGTGGAGTCCATGTAGTCT 2054

Qy       2151 TACCTGTGGAGTTGGAATTCAGACCCGAGATGTGTACTGCCTGCACC------------- 2197
               || ||||| || ||  | |||||| |||| || | ||||    ||||
Db       2055 CACATGTGGGGTCGGCCTACAGACCAGAGACGTCTTCTGCAGCCACCTGCTTTCCAGAGA 2114

Qy       2198 ----CAGGGGACACCCTGCCCCTCCTGAGGAGTGCCGAGATGAAAAGCCCCATGCTTT 2252
                  |       |          ||  ||||  ||| ||   |    ||||||    |  |
Db       2115 GATGAATGAAACAGTCATCCTGGCTGATGAGCTGTGTCGCCAGCCCAAGCCCAGCACGGT 2174

Qy       2253 ACAAGCATGCAATCAGTTTGACTGCCCTCCTGGCTGGCACATTGAAGAATGGCAGCACTC 2312
                ||||| || || |  ||| | ||||| || | |||| ||  || | | ||||||| |||
Db       2175 GCAAGCTTGTAACCGCTTTAATTGCCCCCCAGCCTGGTACCCTGCACAGTGGCAGCCGTG 2234

Qy       2313 TTCCAGGACTTCTGGCGGCGGGAACTCAGAACAGAAGAGTCACCTGTCGGCAGCTGCTAAC 2372
              |||||| || ||||||||||    |||||| |    ||   ||   |||||   |  |
Db       2235 TTCCAGAACGTGTGGCGGGGGTGTTCAGAAACGTGAGGTTCTTTGCAAGCAGCGCATGGC 2294

Qy       2373 GGATGGCAGCTTTTTGAATCTCTCAGATGAATTGTGCCAAGGACCCAAGGCATCGTCTCA 2432
               ||||||||||  || | ||  | ||     || ||  ||   | ||  |  | |  ||
Db       2295 TGATGGCAGCTTCCTGGAGCTTCCTGAGACCTTCTGTTCAGCTTCAAAACCTGCCTGCCA 2354

Qy       2433 CAAGTCCTGTGCCAGGACAGACTGTCCTCCACATTTAGCTGTGGGAGACTGGTCGAAGTG 2492
                |  | ||    |       ||||||||      | |     | |    |||||| |  ||||
Db       2355 GCAAGCATGCAAGAAAGATGACTGTCCCAGCGAGTCGGCTTCTCTCAGACTGGACAGAGTG 2414

Qy       2493 TTCTGTCAGTTGTGGTGTTGGAATCCAGAGAAGAAAGCAGGTGTGTCAAAGGCTGCCAGC 2552
              |||      || || || |  || | ||||||   |||        | || | || | ||
Db       2415 TTCCACAAGCTGCGGGGAAGGCACCCAGACTCGAACCGCCATTTGCCGAAAGATGCTGAA 2474

Qy       2553 CAAAGGTCGGCGCATCCCCCTCAGTGAGATGATGTCCAGGGATCTACCAGGGTTCCCTCT 2612
                |  ||| |        |        ||| |   |   ||||| |    || ||    |   |
Db       2475 AACCGGCCTCTCAACGGTTGTCAATTCCACCCTGTGCCCGCCCCTGCCTTTCTCTTCCTC 2534

Qy       2613 TGTAAGATCTTGCCAGATGCCTGAGTGCAGTAAAATCAAATCAGAGATGAAGACAAAACT 2672
                | ||  | ||   | || |                 || |    ||  |     |     |
Db       2535 CATCAGGCCCTGTATGCTGGC------------AACCTGTGCAAGGCCCCGGCCGCCATC 2582
```

```
Qy       2673 TGGTGAGCAGGGTCCGCAGATCCTCAGTGTCCAGAGAGTCTACATTCAGACAAGGCAAGA 2732
                   |||| | ||||| |||        | |  ||   |||||||| |||||   |      |
Db       2583 CACGAAGCACAGCCCGGCACATCGCCGCCGCAGGAAGGTCTACATACAGACTCGGCAGGCA 2642

Qy       2733 GAAGCGTATTAACCTGACCATTGGTAGCAGAGCCTATTTGCTGCCCAACACATCCGTGAT 2792
               || |     |  || |        | ||  ||    |||||  |||| ||||| ||  | ||| |
Db       2643 GAGGAAGCTGGACTTCGTGGTGGGGGGCTTCGGCTACCTGCTCCCCAAGACGGCGGTGGT 2702

Qy       2793 TATTAAGTGCCCCGTGCGACGATTCCAGAAATCTCTGATCCAGTGGGAGAAGGATGGCCG 2852
                  |     ||||| | |||  |  |||  ||   | ||  |||    |||||||||| ||||
Db       2703 GCTCCGCTGCCGCCGGCGCAGGGTCGGCAAGCCCCTCATCACCTGGGAGAAGGACGGCCA 2762

Qy       2853 TTGCCTCCAGAACTCCAAACCCTTTTTATCACTAATTTAGTTTCAGTAAAAATCCAGGG 2912
                 |||     | ||| |  |   |       |  ||    |   |||    || || ||||||  |
Db       2763 GCACCTCATCAGCTCGACGCACGTCACGCTGCCCCCCTTCGGCTATCTCAAGATCCAGCG 2822

Qy       2913 TCTTGCTGCCCCCGACATCGGCGTGTACCGGTGCATTGCAGGCTCTGCACAGGAAACACT 2972
                ||       || | ||      |||| |||    |||    || ||| | || | |||       |
Db       2823 CCTCAAGCCCTCGGATGCAGGCGTCTACACCTGCTCAGCTGGCCCGGCCCGGGAGCACTT 2882

Qy       2973 TGTGCTCAAGCTCATTGGTACTGAGAACCGGCTCATCGTACGCTC 3017
                |||| | ||||||||| ||        ||   |   |||| | || || ||
Db       2883 TGTGATTAAGCTCATCGGAGGCAACCGCAAGCTCGTGGCCCGGCC 2927
```

```
RESULT 13
US-09-796-753-17
; Sequence 17, Application US/09796753
; Publication No. US20030027998A1
; GENERAL INFORMATION:
;  APPLICANT: McCarthy, Sean A.
;  TITLE OF INVENTION:  SECRETED PROTEINS AND USES THEREOF
;  FILE REFERENCE:  7853-227-999
;  CURRENT APPLICATION NUMBER: US/09/796,753
;  CURRENT FILING DATE:  2001-03-01
;  PRIOR APPLICATION NUMBER:  09/183,175
;  PRIOR FILING DATE:  1998-10-30
;  PRIOR APPLICATION NUMBER:  09/223,094
;  PRIOR FILING DATE:  1998-12-30
;  PRIOR APPLICATION NUMBER:  09/223,546
;  PRIOR FILING DATE:  1998-12-30
;  PRIOR APPLICATION NUMBER:  09/224,246
;  PRIOR FILING DATE:  1998-12-30
;  PRIOR APPLICATION NUMBER:  09/259,388
;  PRIOR FILING DATE:  1999-02-26
;  PRIOR APPLICATION NUMBER:  60/122,458
;  PRIOR FILING DATE:  1999-03-01
;  PRIOR APPLICATION NUMBER:  09/312,359
;  PRIOR FILING DATE:  1999-05-14
;  PRIOR APPLICATION NUMBER:  09/336,536
;  PRIOR FILING DATE:  1999-06-18
;  PRIOR APPLICATION NUMBER:  09/342,687
;  PRIOR FILING DATE:  1999-06-29
;  PRIOR APPLICATION NUMBER:  09/345,464
;  PRIOR FILING DATE:  1999-06-30
;  PRIOR APPLICATION NUMBER:  09/365,164
;  PRIOR FILING DATE:  1999-07-30
;  PRIOR APPLICATION NUMBER:  09/399,723
;  PRIOR FILING DATE:  1999-09-20
;  PRIOR APPLICATION NUMBER:  09/409,634
;  PRIOR FILING DATE:  1999-09-30
;  PRIOR APPLICATION NUMBER:  09/471,179
;  PRIOR FILING DATE:  1999-12-23
;  PRIOR APPLICATION NUMBER:  09/474,071
;  PRIOR FILING DATE:  1999-12-29
;  PRIOR APPLICATION NUMBER:  09/474,072
;  PRIOR FILING DATE:  1999-12-29
;  PRIOR APPLICATION NUMBER:  09/514,010
;  PRIOR FILING DATE:  2000-02-25
;  PRIOR APPLICATION NUMBER:  09/516,745
;  PRIOR FILING DATE:  2000-03-01
;  PRIOR APPLICATION NUMBER:  09/572,002
;  PRIOR FILING DATE:  2000-05-14
;    PRIOR APPLICATION NUMBER: 09/597,993
;    PRIOR FILING DATE: 2000-06-19
;    PRIOR APPLICATION NUMBER: 09/599,596
;    PRIOR FILING DATE: 2000-06-22
;    PRIOR APPLICATION NUMBER: 09/630,334
;    PRIOR FILING DATE: 2000-07-31
;    PRIOR APPLICATION NUMBER: 09/606,565
;    PRIOR FILING DATE: 2000-06-29
;    PRIOR APPLICATION NUMBER: 09/606,317
;    PRIOR FILING DATE: 2000-06-29
;    PRIOR APPLICATION NUMBER: 09/665,666
;    PRIOR FILING DATE: 2000-09-20
;    PRIOR APPLICATION NUMBER: 09/677,751
;    PRIOR FILING DATE: 2000-09-30
;    NUMBER OF SEQ ID NOS: 162
;  SEQ ID NO 17
;     LENGTH: 2689
;     TYPE: DNA
;     ORGANISM: Homo sapiens
;     FEATURE:
;     NAME/KEY: CDS
;     LOCATION: (67)...(1449)
US-09-796-753-17

  Query Match             18.0%;  Score 913;  DB 9;  Length 2689;
  Best Local Similarity   62.5%;  Pred. No. 5.1e-263;
  Matches 1512;  Conservative    0;  Mismatches  850;  Indels   59;  Gaps    3;

Qy        186 CCAGACCTCAAGAAACACTCGTTCAGATGAAGACAAAGATGGCAACTGGGATGCTTGGGG 245
                || ||  || || | | | :: :| || || |||    || |||    ||||:||| ||||
Db        126 CCTGATTTCAGGACCGCACTGTCCGAGGAGGACCGGGACGGCTGTGGGATGCCTGGGG 185

Qy        246 CGACTGGAGGACTGCTCCTGGACCTGTGGGGGAGGAGCATCATATTCTTGCGGAGATG 305
                |    ||||||| ||||| |: ||||| ||||||| || || || |||||| || | ||
Db        186 CCCATGGAGTGAATGCTCAGCACCTGCGGGGAGGGGCCTCCTACTCTCTGAGGCGCTG 245

Qy        306 TTTGACTGGAAGGAATCCGGAAGGGAGAACATTCGGTACAAGACATGCAGCAATCATGA 365
                |||    | | | ||   ||||||||       || || || ||||  ||||||| |||    ||
Db        246 CCTGAGCAGCAAGAGCTGTGAAGGAAGAAATATCCGATACAGAACATGCAGTAATGTGGA 305

Qy        366 CTGCCCTCCAGATGCAGAAGACTTCAGAGCCCAGCAGTGCTCAGCCTACAATGATGTCCA 425
                || | | ||||: ||| :||||: ||| ||||| |||||.|    | .|| | ||  .
Db        306 CTGCCCACCAGAAGCAGGTGATTTCGAGCTCAGCAATGCTCAGCTCATAATGATGTCAA 365

Qy        426 GTATGAGGGGATTACTATGAATGGATTCCAGGATATAAGGATCTGCTGCCCCGTGCGC 485
                . | | | |       .  .:||||| |      | |||||| ||||      | :
Db        366 GTACCATGGCCAGTTTTATGAATGGCTTCCTGTGTCTAATGACCCTGACAACCCATGTTC 425

Qy        486 ACTCAAGTGTCATGCACAAGGACAAAATTTGGTGGTGGAGCTGGCACCTAAGGTACTGGA 545
                || |.||| |.  |  .  .  |: : |||| || || || |.|||||:|:||    :
Db        426 ATCCAAGTGCCAAGCCAAAGGAACAACCTTGGTTGTTGAACTAGCACCTAAGGTCTTAGA 485

Qy        546 TGGAACTCGTTGCAACACGGACTCCTTGGACATGTGTATCAGTGGCATCTGTCAGGCAGT 605
                ||| || |||||| | : : || ||||| |||||: ||||||||   :  |  .|    |:
Db        486 TGGTACGCGTTGCTATACAGAATCTTTGGATATGTGCATCAGTGGTTTATGCCAAATTGT 545

Qy        606 GGGCTGCGATCGGCAACTGGGAAGCAATGCCAAGGAGGACAACTGTGGAGTCTGTGCCGG 665
                :|||||||||  || ||:|||||||||   | ||||||| || |||||||| |||||   :||
Db        546 TGGCTGCGATCACCAGCTGGGAAGCACCGTCAAGGAAGATAACTGTGGGGTCTGCAACGG 605

Qy        666 CGATGGCTCCACCTGCAGGCTTGTTACGGGGACAATCAAAGTCACACGTTCTCTCCTGAAAA 725
                ||| | ||||||||| |||| || || || || |  || || || |  | ||  |    ||
Db        606 AGATGGGTCCACCTGCCGGCTGGTCCGAGGGCAGTATAAATCCCAGCTCTCCGCAACCAA 665

Qy        726 AAGAGAAGAAAATGTAATCGCTGTTCCTTTGGGAAGTCGAAGTGTGAGAATTATAGTGAA 785
                .  .| |. | ||   |. :   |||| |  ||||| ||  | |  |  :|    | :|
Db        666 ATCGGATGATACTGTGGTTGCAATTCCCTATGGAAGTAGACATATTCGCCTTGTCTTAAA 725

Qy        786 AGGACCTGCCCACCTCTTTATTGAATCAAAAACACTTCAAGGAAGCAAAGGAGAACACAG 845
                |:| :  |   ||| | |  .  . |.| | ||||| :| || |  |  ||||| ||| .|||
Db        726 AGGTCCTGATCACTTATATCTGGAAACCAAAACCCTCCAGGGGACTAAAGGTGAAAACAG 785

Qy        846 CTTTAACAGCCCCGGCGTCTTTGTCGTAGAAAACACAACAGTGGAATTTCAGAGGGGCTC 905
                | | |  | | ||   ||  | || || ||  | |  |||| :| ||||       |
Db        786 TCTCAGCTCCACAGGAACTTTCCTTGTGGACAATTCTAGTGTGGACTTCCAGAAATTTCC 845

Qy        906 CGAGAGGCAAACTTTTAAGATTCCAGGACCTCTGATGGCTGATTTCATCTTCAAGACCAG 965
                || |   | |   | |  ||  | ||||| || |  || ||||||||  |||||| |
Db        846 AGACAAAGAGATACTGAGAATGGCTGGACCACTCACAGCAGATTTCATTGTCAAGATTCG 905

Qy        966 GTACACTGCAGCCAAAGACAGCGTGGTTCAGTTCTTCTTTTACCAGCCCATCAGTCATCA 1025
                || | :   ||   |||||    || :|||||  |||: || || :|: | |  ||| |
Db        906 TAACTCGGGCTCCGCTGACAGTACAGTCCAGTTCATCTTCTATCAACCCATCATCCACCG 965

Qy       1026 GTGGAGACAAACTGACTTCTTTCCCTGCACTGTGACCTGTGGAGGAGGTTATCAGCTCAA 1085
                |:|||   | || || ||||||| ||| |  |   |: ||||. ||||| | ||||| |
Db        966 ATGGAGGGAGACGGATTTCTTTCCTTGCTCAGCAACCTGTGGAGGAGGTTATCAGCTGAC 1025

Qy       1086 TTCTGCTGAATGTGTGGATATCCGCTTGAAGAGGGTAGTTCCTGACCATTATTGTCACTA 1145
                 :| ||||| ||      ||| |  |      ||  | || || |  |||| |  ||||||||
Db       1026 ATCGGCTGAGTGCTACGATCTGAGGAGCAACCGTGTGGTTGCTGACCAATACTGTCACTA 1085

Qy       1146 CTACCCTGAAAATGTAAAACCAAAACCAAAACTGAAGGAATGCAGCATGGATCCCTGCCC 1205
                ||||| :| ||  | ||||| :|||| || ||  ||| || | | |||:|:| |  ||
Db       1086 TTACCCAGAGAACATCAAACCCAAACCCAAGCTTCAGGAGTGCAACTTGGATCCTTGTCC 1145

Qy       1206 ATC-------------------------AAGTGATGGATTTAAAGAGATAATGCCCTAT 1239
                | |                         ||||| ||||  |:  |:|| ||||| |||
Db       1146 AGCCAGAGGCTTGGCATTATTGTTCCAACAGTGACGGATACAACCAGATCATGCCTTAT 1205

Qy       1240 GACCACTTCCAACCTCTTCCTCGCTGGGAACATAATCCTTGGACTGCATGTTCCGTGTCC 1299
                |||| || ||| || |||||||| |||||      |  || ||||| || || |||    ||
Db       1206 GACCTCTACCATCCCCTTCCTCGGTGGGAGGCCACCCCATGGACCGCGTGCTCCTCCTCG 1265

Qy       1300 TGTGGAGGAGGGATTCAGAGACGGAGCTTTGTGTGTGTAGAGGAATCCATGCATGGAGAG 1359
                ||||| || || || ||| : |||     ||    ||||| |||||    ||| || ||   |
Db       1266 TGTGGGGGGGGCATCCAGAGCCGGGCAGTTTCCTGTGTGGAGGAGGACATCCAGGGGCAT 1325

Qy       1360 ATATTGCAGGTGGAAGAATGGAAGTGCATGTACCCACCCAAACCCAAGGTTATGCAAACT 1419
                |          ||||||||   | || |||||||| | || ||               |   |||   |
Db       1326 GTCACTTCAGTGGAAGAGTGGAAATGCATGTACACCCCTAAGATGCCCATCGCGCAGCCC 1385

Qy       1420 TGTAATCTGTTTGATTGCCCCAAGTGGATTGGCATGGAGTGGTCTCAGTGCACAGTGACT 1479
                || ||  | ||||| |||:| || :|| : ||   :||| |:|:|| ||||  ||||||
Db       1386 TGCAACATTTTTGACTGCCCTAAATGGCTGGCACAGGAGTGGTCTCCGTGCACAGTGACA 1445

Qy       1480 TGTGGCCGAGGGTTACGGTACCGGGTTGTTCTGTGTATTAACCACCGCGGAGAGCATGTT 1539
                |||||||   ||   |   | :| || || |: || :| |:   :||| || |||   |||
Db       1446 TGTGGCCAGGGCCTCAGATACCGTGTGGTCCTCTGCATCGACCATCGAGGAATGCACACA 1505

Qy       1540 GGGGGCTGCAATCCACAACTGAAGTTACACATCAAAGAAGAATGTGTCATTCCCATCCCG 1599
                || ||||| |   ||| ||    |||     |||||| ||||| :|||:   || | ||||   ||
Db       1506 GGAGGCTGTAGCCCAAAAACAAAGCCCCACATAAAAGAGGAATGCATCGTACCCACTCCC 1565

Qy       1600 TGTTATAAACCAAAAGAAAAAAGTCCAGTGGAAGCAAAATTGCCTTGGCTGAAACAAGCA 1659
                || ||||||||| ||||| |||   |||||| || || || ||||| ||| | ||||||||
Db       1566 TGCTATAAACCCAAAGAGAAACTTCCAGTCGAGGCCAAGTTGCCATGGTTCAAACAAGCT 1625

Qy       1660 CAAGAACTAGAAGAGACCAGAATAGCAACAGAAGAACCAACGTTCATTCCAGAACCCTGG 1719
                ||||| ||||||||                |      |||| || ||   |||||| ||||| |||||
Db       1626 CAAGAGCTAGAAGAAGGAGCTGCTGTGTCAGAGGAGCCCTCGTTCATCCCAGAGGCCTGG 1685

Qy       1720 TCAGCCTGCAGTACCACGTGTGGGCCAGGTGTGCAGGTCCGCGAGGTGAAGTGCCGTGTG 1779
                || |||||||     ||| |:||     ||       ||||| ||      || | ||||| |||
Db       1686 TCGGCCTGCACAGTCACCTGTGGTGTGGGGACCCAGGTGCGAATAGTCAGGTGCCAGGTG 1745

Qy       1780 CTCCTCACATTCACGCAGACTGAGACTGAGCTGCCCGAGGAAGAGTGTGAAGGCCCCAAG 1839
                |||||   | ||:   |||     | | |||| ||||:      |: |||||||||||| |:||||
Db       1746 CTCCTGTCTTTCTCTCAGTCCGTGGCTGACCTGCCTATTGACGAGTGTGAAGGCCCCAAG 1805

Qy       1840 CTGCCCACC-----------------------GAACGGCCTGCCTCCTGGAAGCATGTGATGAGAGC 1884
                |   |  ||                        | | | | :|||  |  ||||      ||    ||
Db       1806 CCAGCATCCCAGCGTGCCTGTTATGCAGGCCCATGCAGCGGGAAATTCCTCAGTTCAAC 1865

Qy       1885 CCGGCCTCCCGAGAGCTAGACATCCCTCTCCCTGAGGACAGTGAGACGACTTACGACTGG 1944
                || | |     |||        | |   |  ||   ||||    ||    |    || ||||||
Db   1866 CCAGACGAGACAGATGGGCTCTTTGCTGCCCTGCACCATTTCCACCAGCTCTATGACTCC 1925

Qy   1945 GAGTACGCTGGGTTCACCCCTTGCACAGCAACATGCTTGGCAGGCCATCAAGAAGCCATA 2004
          ||||| |  |||||||||   ||| | |   | ||    |||||    || || ||  |
Db   1926 GAGTATGAGGGGTTCACCAAGTGCTCCGAGTCCTGTGGAGGAGGTGTCCAGGAGGCTGTG 1985

Qy   2005 GCAGTGTGCTTACATATCCAGACCCAGCAGACACTCAATGACAGCTTGTGTCATATCCTC 2064
          |     |||||  | |  ||||| | | || | |   | || | | |||| |  |    |
Db   1986 GTCACCTCCTTTAACAAATACACTCGGAGCCTGCTTAGGAGAACCTCCCCCCTGACCAGC 2045

Qy   2065 CACCGTCCTCCAGCCATGAGCCAGGCCTGTAACACACACCCCTCTCCCCCAGCCTGCCAT 2124
          | ||| || |||    |     || |||| ||    || ||||| ||  | ||||||  |
Db   2046 CGCCGGCCCCCACAGCTCCTGAAGTCCTGCAATTTGGATCCCTCCCCAGCAAGCTGCCAA 2105

Qy   2125 GTGGGCTCTTGGGGGCCCTGCTCAGCTACCTGTGGAGTTGGAATTCAGACCAGACATTTC 2184
           | |||   ||| | || ||       || ||||| || ||  | |||||| |||| ||
Db   2106 ATTGGCAAGTGGAGTCCATCTAGTCTCACATGTGGGTCCGGCCTACAGACCAGAGACGTC 2165

Qy   2185 TACTGCCTGCACC------------------CAGGCCAGACCCCTGCCCCTCCTCAGGAG 2226
          | ||||   ||||                    |       |       ||  ||||  |
Db   2166 TTCTGCAGCCACCTGCTTTCCAGAGAGATCAATGAAACAGTCATCCTGGCTGATGACCTG 2225

Qy   2227 TCCCCAGATTAAAACCCCCATCTTTACAATCATGCAATTACTTTCAGTGCCCCTGCTGCC 2286
          || ||  |    |||||     |  | ||||| ||  | |   || | ||||| || | |
Db   2226 TGTCCCAGCTCAAGCCCAGCACGGTGTAAGCTTGTAATTGCTTTAATTCCGCCCCCAGCC 2285

Qy   2287 TGCCACATTGAACAATGGCAGCAGTGTTCCAGGACTTCTCCGGGGGGAACTCAGAATAGA 2346
          ||| ||  || | | ||||||| ||||||||| || |||||||||||   ||||||  |
Db   2286 TCGTACCCTCCACAGTGCCAGCCGTGTTCTACAACGTCTGGCGGCCCTGTTCATAAATCT 2345

Qy   2347 AGACTCACCTGTCGGCAGCTGCTAACGGATGGCAGCTTTTTGAATCTCTCAGATGAATTG 2406
             ||    ||   |||||   |  | |||||||||||  || | ||  | ||    ||
Db   2346 GAGGTTCTTTGCAAGCAGGCATGGCTGATGGCAGCTTCCTGGAGCTTCCTGAGACCTTTC 2405

Qy   2407 TGCCAAGGACCCAAGGCATCGTCTCACAAGTCCTGTGCCAGGACAGACTGTCCCTCACAT 2466
          ||   ||   | ||  |  | |  ||  |  | ||    |     ||||||||     |
Db   2406 TGTTCAGCTTCAAAACCTGCCTGCCAGCAAGCATGCAAGAAAGATGACTGTCCCAGCCGAG 2465

Qy   2467 TTAGCTGTGGGAGACTGGTCGAAGTGTTCTGTCAGTTGTGGTGTGGAATCCAGAGAAGA 2526
          |    | |   |||||||| | |||||||    || || || |  || | |||||   ||
Db   2466 TGGCTTCTCTCAGACTGGACAGAGTGTTCCACAAGCTGCGGCCAAGCCACCCAGACTCCA 2525

Qy   2527 AAGCAGGTGTGTCAAAGGCTG 2547
          |      | || | || | ||
Db   2526 AGCGCCATTTGCCGAAAGATG 2546

RESULT 14
US-09-796-753-19
; Sequence 19, Application US/09796753
; Publication No. US20030027998A1
; GENERAL INFORMATION:
;  APPLICANT: McCarthy, Sean A.
;  TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
;  FILE REFERENCE: 7853-227-999
;  CURRENT APPLICATION NUMBER: US/09/796,753
;  CURRENT FILING DATE: 2001-03-01
;  PRIOR APPLICATION NUMBER: 09/183,175
;  PRIOR FILING DATE: 1998-10-30
;  PRIOR APPLICATION NUMBER: 09/223,094
;  PRIOR FILING DATE: 1998-12-30
;  PRIOR APPLICATION NUMBER: 09/223,546
;  PRIOR FILING DATE: 1998-12-30
;  PRIOR APPLICATION NUMBER: 09/224,246
;  PRIOR FILING DATE: 1998-12-30
;  PRIOR APPLICATION NUMBER: 09/259,388
;  PRIOR FILING DATE: 1999-02-26
;  PRIOR APPLICATION NUMBER: 60/122,458
;  PRIOR FILING DATE: 1999-03-01
;  PRIOR APPLICATION NUMBER: 09/312,359
;  PRIOR FILING DATE: 1999-05-14
;  PRIOR APPLICATION NUMBER: 09/336,536
;  PRIOR FILING DATE: 1999-06-18
;  PRIOR APPLICATION NUMBER: 09/342,687
;  PRIOR FILING DATE: 1999-06-29
;  PRIOR APPLICATION NUMBER: 09/345,464
;  PRIOR FILING DATE: 1999-06-30
;  PRIOR APPLICATION NUMBER: 09/365,164
;  PRIOR FILING DATE: 1999-07-30
;  PRIOR APPLICATION NUMBER: 09/399,723
;  PRIOR FILING DATE: 1999-09-20
;  PRIOR APPLICATION NUMBER: 09/409,634
;  PRIOR FILING DATE: 1999-09-30
;  PRIOR APPLICATION NUMBER: 09/471,179
;  PRIOR FILING DATE: 1999-12-23
;  PRIOR APPLICATION NUMBER: 09/474,071
;  PRIOR FILING DATE: 1999-12-29
;  PRIOR APPLICATION NUMBER: 09/474,072
;  PRIOR FILING DATE: 1999-12-29
;  PRIOR APPLICATION NUMBER: 09/514,010
;  PRIOR FILING DATE: 2000-02-25
;  PRIOR APPLICATION NUMBER: 09/516,745
;  PRIOR FILING DATE: 2000-03-01
;  PRIOR APPLICATION NUMBER: 09/572,092
;  PRIOR FILING DATE: 2000-05-14
;  PRIOR APPLICATION NUMBER: 09/597,993
;  PRIOR FILING DATE: 2000-06-19
;  PRIOR APPLICATION NUMBER: 09/599,596
;  PRIOR FILING DATE: 2000-06-22
;  PRIOR APPLICATION NUMBER: 09/629,334
;  PRIOR FILING DATE: 2000-07-31
;  PRIOR APPLICATION NUMBER: 09/606,565
;  PRIOR FILING DATE: 2000-06-29
;  PRIOR APPLICATION NUMBER: 09/606,317
;  PRIOR FILING DATE: 2000-06-29
;  PRIOR APPLICATION NUMBER: 09/665,666
;  PRIOR FILING DATE: 2000-09-20
;  PRIOR APPLICATION NUMBER: 09/677,751
;  PRIOR FILING DATE: 2000-09-30
;  NUMBER OF SEQ ID NOS: 162
; SEQ ID NO 19
;   LENGTH: 2689
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (67)...(2689)
US-09-796-753-19

  Query Match             18.0%;  Score 912;  DB 9;  Length 2689;
  Best Local Similarity   62.4%;  Pred. No. 1e-262;
  Matches 1512;  Conservative    0;  Mismatches  850;  Indels   60;  Gaps    3;

Qy    186 CCAGACCTCAAGAAACACTCGTTCAGATGAAGACAAAGATGGCAACTGGGATGCTTGGGG 245
          || ||  || || | | | || || || || |||   || |||   |||||||| |||||
Db    126 CCTGAGTTCCAGGACCGCACGCTCCGAGGAGGACCGGGACGGCCTGTGGGATGCCTGGGG 185

Qy    246 CGACTGGAGTGACTGCTCCCGGACCTGTGGGGGAGGAGCATCATATTCTCTGCGGAGATG 305
          |   |||||||| ||||| || ||||| |||||||| || || || |||||| || | ||
Db    186 CCCATGGAGTGAATGCTCACGCACCTGCGGGGGAGGGGCCTCCTACTCTCTGAGGCGCTG 245

Qy    306 TTTGACTGGAAGGAATTGTGAAGGGCAGAACATTCGGTACAAGACATGCAGCAATCATGA 365
            |||   | | ||  ||||||||    || || || ||||  |||||||| |||   ||
Db    246 CCTGAGCAGCAAGAGCTGTGAAGGAAGAAATATCCGATACAGAACATGCAGTAATGTGGA 305

Qy    366 CTGCCCTCCAGATGCAGAAGATTTCAGAGCCCAGCAGTGCTCAGCCTACAATGATGTCCA 425
          |||||| ||||| ||||  |||||| |||| ||||| ||||||||  | ||||||||| |
Db    306 CTGCCCACCAGAAGCAGGTGATTTCCGAGCTCAGCAATGCTCAGCTCATAATGATGTCAA 365

Qy    426 GTATCAGGGGCATTACTATGAATGGCTTCCACGATATAATGATCCTGCTGCCCCGTGTGC 485
          | | || || || |  ||||||||||||||    | |||||| ||||    ||| ||| |
Db    366 GCACCATGGCCAGTTTTATGAATGGCTTCCTGTGTCTAATGACCCTGACAACCCATGTTC 425

Qy    486 ACTCAAGTGTCATGCACAAGGACAAAACTTGGTGGTGGAGCTGGCACCTAAGGTACTGGA 545
          ||||||||| || ||  |||||  || | |||| || || || |||||||||||  | ||
Db    426 ACTCAAGTGCCAAGCCAAAGGAACAACCCTGGTTGTTGAACTAGCACCTAAGGTCTTAGA 485
```

```
Qy    546 TCGAACTCGTCGCAACACGGACTCGCTTGGACATGTGTATCAGTGGCATCTGTCAGGCAGT 605
Db    486 TGGTACGGCGTTGCTATACAGAATCTTTGGATATGTGCATCAGTGGTTTATGCCAAATTGT 545

Qy    606 GGGCTGGGATCGGCAACTGGGAAGCAATGCCAAGGAGGACAACTGTGGAGTCTGTGCCGG 665
Db    546 TGGCTGGGATCACCAGCTGGGAAGCACCGTCAAGGAAGATAACTGTGGGGTCTGCAAGGG 605

Qy    666 CGATGGCTCCACCTGCATCCTTGGTACGGGGACAATCAAAGTCACAGGTTTCTCCTGAAAA 725
Db    606 AGATGGGTCCACCTGCCGGCTGGTCGAGGGCAGTATAAATCCCAGCTCTCCGCAACCAA 665

Qy    726 AAGAGAAGAAAATGTAATTGGTGGTTCCTTTGGGAAGTCGAAGTGTGAGAATTACAGTGAA 785
Db    666 ATCGGATGATACTGTGGTTGAATTCCCTATGGAAGTAGACATATTCGCCTTGTCTTAAA 725

Qy    786 AGGACCTGCCCACCTGTTTATTGAATCAAAAACACTCCAAGGAAGTAAATGAGAACACAG 845
Db    726 AGGTCCTGATCACTTATATCTGGAAACCAAAACCCTCCAGGGGACTAAAGGTGAAAACAG 785

Qy    846 CTTTAACAGCCCCGGCGACTTTGTCGTAGAAAACACAACAGTGGAATTTCAGAGGGTGC 905
Db    786 TCTCAGCTCCACAGGAACTTTCTTGTGGACAATTCTAGTGTGGACTTCCAGAAATTTCC 845

Qy    906 CCAGAGGCAAACTTTTAAGATTCCAGGACCTCTGATGGCTGATTCATCTTCAAGACTAG 965
Db    846 AGACAAAGAGATACTGAGAATGGCTGGACCACTCACAGCAGATTTCATTGTCAAGATTCG 905

Qy    966 GTACACTGGAGCCAAAGACAGCGTGGTTCAGTTCTTCTTTACCAGCCCATCAGTCATCA 1025
Db    906 TAACTCGGGCTCCCTGACAGTACAGTCCAGTTCATCTTCTATCAACCCATCATCACCG 965

Qy   1026 GTGGAGACAAACTGACTTCTTTCCCTGCACTGTGACGTGTGGAGGAGGTTATTAGCTCAA 1085
Db    966 ATGGAGGGACAGGCATTTCTTTCCTTGCTCAGCAACCTGTGGAGGAGGTTATCAGCTGAC 1025

Qy   1086 TTCTGCTGAATGTTGGGATATCCGCTTGAAGAGGGTAGTTCCTGACCATTATTGTCACTA 1145
Db   1026 ATCGGCTGAGTGCTACGATCTGAGGAGCAACCGTGTGGTTGCTGACCAATACTGTCACTA 1085

Qy   1146 CTACCCTGAAAATGTAAAACCAAAACCAAAACTGAAGGAATGCAGCATGGATCCCTGCCC 1205
Db   1086 TTACCCAGAGAACATCAAACCCAAACCCAAGCTTCAGGAGTGCAACTTGGATCCTTGTCC 1145

Qy   1206 ATC                          AAGTGATGGATTTAAAGAGATAATGCCCTA 1238
Db   1146 AGCCCAGAGGGTTGGCATTATTGTTCCTAACAGTGACGGATACAAGCAGATCATGCCCTA 1205

Qy   1239 TGACCACTTCCAACCTCTTCCTCGCTGGGAACATAATCCTTGGACTGCATGTTCCGTGTC 1298
Db   1206 TGACCTCTACCATCCCCTTCCTCGGTGGGAGGCCACCCCATGGACCGCGTGCTCCTCCTC 1265

Qy   1299 CTGTGGAGGAGGGATTCAGAGACGGAGCTTTGTGTGTGTAGAGGAATCCATGCATGGAGA 1358
Db   1266 GTGTGGGGGGGGCATCCAGAGCCGGGCAGTTTCCTGTGTGGAGGAGGACATCCAGGGGCA 1325

Qy   1359 GATATTGCAGGTGGAAGAATGGAAGTGCATGTATGGCACCCAAACCCAAGGTTATGCAAAC 1418
Db   1326 TGTCACTTCAGTGGAAGAGTGGAAATGCATGTACACCCCTAAGATGCCCATCGCGCAGCC 1385

Qy   1419 TTGTAATCTGTTTGATTGCCCCAAGTGGATTGCCATGGAGTGGTCTCAGTGCACAGTGAC 1478
Db   1386 CTGCAACATTTTTGACTGCCCTAAATGGCTGGCACAGGAGTGGTCTCCGTGCACAGTGAC 1445

Qy   1479 TTGTGGCCGAGGGTTACGGTACCGGGTTGTTCTGTGTATTAACCACCGGGGAGAGCATGT 1538
Db   1446 ATGTGGCCAGGGCCTCAGATACCGTGTGGTCCTTTGCATCGACCATCGAGGAATGCACAC 1505

Qy   1539 TGGGGGCTGCAATCCACAACTGAAGTTACACATCAAAGAAGAATGTGTCATTCCCATCCC 1598
Db   1506 AGGAGGCTGTAGCCCAAAAACAAAGCCCCACATAAAAGAGGAATGCATCGTACCCACTCC 1565

Qy   1599 GTGTTATAAACCAAAAGAAAAAGTCCAGTGGAAGCAAAATTGCCTTGCCTGAAACAAGC 1658
Db   1566 CTGCTATAAACCCAAAGAGAAACTTCCAGTCGAGGCCAAGTTGCCATGGTTCAAACAAGC 1625

Qy   1659 ACAAGAACTAGAAGAGACCAGAATAGCAACAGAAGAACCAACGTTCATTCCAGAACCCTG 1718
Db   1626 TCAAGAGCTAGAAGAAGGAGCTGCTGTGTCAGAGGAGCCCTCGTTCATCCCAGAGGCCTG 1685

Qy   1719 GTCAGCCTGCAGTACCACGTGTGGGCCAGGTGTGCAGGTCCGCGAGGTGAAGTGCCGTGT 1778
Db   1686 GTCGGCCTGCACAGTCACCTGTGGTGTGGGGACCCAGGTGCGAATAGTCAGGTGCCAGGT 1745

Qy   1779 GCTCCTCACATTCACGCAGACTGAGACTGAGCTGCCTGAGGAAGAGTGTGAAGGCCCCAA 1838
Db   1746 GCTCCTGTCTTCTCTCAGTCCCGTGGCTGACCTGCCTATTGACGAGTGTGAAGGGCCCAA 1805

Qy   1839 GCTGCCCACC                GAACGGCCCTGCCTCCTGGAAGCATGTGATGAGAG 1883
Db   1806 GCCAGCATCCCAGCCTGCCTGTTATGCAGGCCATGCAGCGGGAAATTCCTGAGTTCAA 1865

Qy   1884 CCCGGCCTCCGAGAGCTAGACATCCCTCTTCCTGAGGAGAGGAGACGACTTACGACTG 1943
Db   1866 CCCAGACGAGACAGATGGCTCTTTGGTGGCTGCAGGATTTGACGAGCTATGACTG 1925

Qy   1944 GGAGTACGCTGGCTTCACCCCTTGTACAGCAACATGCTTGGGAGGCCATCAAGAAGCCAT 2003
Db   1926 GGAGTATGAGGGGTTCACCAAGTGCTCCGAGTCCTGTGGAGGAGGTGTCCAGGAGGCTGT 1985

Qy   2004 AGCAGTGTGCTTACATATCCAGACCAGCAGACAGTCAAGGACAGCTTGTGTGATATGGT 2063
Db   1986 GGTGAGCTGCTTGAACAAACAGACTGGGAGCCTGCTGAGGAGAACCTGCTGTGACCAG 2045

Qy   2064 CCACCGTCCTCCAGCCATGAGCCAGGCCTGTAACACAGAGCCCTGTCCCCCCAGGTGGCA 2123
Db   2046 CCGCCGGCCCCCCACAGCTCCTGAAGTCCTGCAATTGGATCCCTGCCCAGCAAGGTGGGA 2105

Qy   2124 TGTGGGCTCTTGGGGGGCCCTGCTCAGCTACCTGTGGAGTTGGAATCCAGACCTGAGATGT 2183
Db   2106 AATTGGCAAGTGGAGTCCATGTAGTCTCACATGTGGGCTGGGCTACAGACCAGAGACGT 2165

Qy   2184 GTACTGCCTGCACC                    CAGGGAGACCCTGCCCTCCTGAGGA 2225
Db   2166 CTTCTGCAGCCACCTGCTTTCAGAGAGATGAATGAAACAGTCATCCTGGCTGATGAGCT 2225

Qy   2226 GTGCCGAGATGAAAAGCCCCATGTTTTACAAGCATGCAATCAGTTTGACTGCCTTCCTGG 2285
Db   2226 GTGTCGCCAGCCCAAGCCCAGCACGGTGCAAGCTTGTAACCGGCTTAATTGCCTCCCAGC 2285

Qy   2286 CTGGCACATTGAAGAATGGCAGCAGTGTTCTAGGACTTGTGGCGGGGGAACTCAGAACAG 2345
Db   2286 CTGGTACCCTGCACAGTGGCAGCCGTGTTCCAGAACGTGTGGCGGGGGTGTTCAGAAACG 2345

Qy   2346 AAGAGTCACCTGTCGGCAGCTGCTAACGGATGGCAGCTTTTTCAATCTCTCAGATGAATT 2405
Db   2346 TGAGGTTCTTTGCAAGCAGCGCATGGCTGATGGCAGCTTCCTCGAGCTTCCTGAGACCTT 2405

Qy   2406 GTGCCAAGGACCCAAGGCATCGTCTCACAAGTCCTGTGCCAGGACAGACTGTCCTCCACA 2465
Db   2406 CTGTTCAGCTTCAAAACCTGCCTGCCAGCAAGATGCAAGAAAGATGACTGTCCCAGCGA 2465

Qy   2466 TTTAGCTGTTGGGAGACTGGTTGAAGTGTTCTGTTACTTGTTGTTGGAATCCAGAGAAG 2525
Db   2466 GTGGCTTCTCTCAGACTGGAAGAGTGTTCACAAGCTGGGGGAAGTACCTAGACTCG 2525

Qy   2526 AAAGCAGGTGTGTCAAAGCCTG 2547
Db   2526 AAGCGCCATTTGCCGAAAGATG 2547


RESULT 15
US-09-800-729-53
; Sequence 53, Application US/09800729
; Patent No. US20020068319A1
; GENERAL INFORMATION:
;  APPLICANT:  Ni et al.
;  TITLE OF INVENTION:  32 Human secreted proteins
;  FILE REFERENCE:  PZ044P1
;  CURRENT APPLICATION NUMBER: US/09/800,729
;  CURRENT FILING DATE:  2001-03-08
;  PRIOR APPLICATION NUMBER:  PCT/US00/26013
;  PRIOR FILING DATE:  2000-09-22
;  PRIOR APPLICATION NUMBER:  60/155,709
;  PRIOR FILING DATE:  1999-09-24
;  NUMBER OF SEQ ID NOS:  217
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 53
;   LENGTH: 3559
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (3363)
;   OTHER INFORMATION: n equals a,t,g, or c
US-09-800-729-53

  Query Match             16.6%;  Score 843.2;  DB 10;  Length 3559;
  Best Local Similarity   58.6%;  Pred. No. 5.9e-242;
  Matches 1584;  Conservative    7;  Mismatches 1015;  Indels   96;  Gaps    4;

Qy        349 ACATGCAGCAATCATGACTGCCCTCCAGATGCAGAAGATTTCAGAGCCCAGCAGTGCTCA 408
              |||||||||  |   || ||||| ||| | ||::  |||||| |||| ||||| ||||||
Db         19 ACATGCAGCCCTTTGGAATGCCCACCAAAAGCMSGTGATTTCCGAGCTCAGCAATGCTCA 78

Qy        409 GCCTACAATGATGTCCAGTATCAGGGGCATTACTATGAATGGCTTCCACGATATAATGAT 468
              ||  | ||||||||| || | || || || |  ||||||||||||||    | ||||||
Db         79 GCTCATAATGATGTCAAGCACCATGGCCAGTTTTATGAATGGCTTCCTGTGTCTAATGAC 138

Qy        469 CCTGCTGCCCCGTGTGCACTCAAGTGTCATGCACAAGGACAAAACTTGGTGGTGGAGCTG 528
              ||||    ||| ||| |||||||||| || ||  |||||  || | |||| || || ||
Db        139 CCTGACAACCCATGTTCACTCAAGTGCCAAGCCAAAGGAACAACCCTGGTTGTTGAACTA 198

Qy        529 GCACCTAAGGTACTGGATGGAACTCGTTGCAACACGGACTCCTTGGACATGTGTATCAGT 588
              |||||||||||  | ||||| || |||||| | || || || ||||| ||||| ||||||
Db        199 GCACCTAAGGTCTTAGATGGTACGCGTTGCTATACAGAATCTTTGGATATGTGCATCAGT 258

Qy        589 GGCATCTGTCAGGCAGTGGGCTGCGATCGGCAACTGGGAAGCAATGCCAAGGAGGACAAC 648
              ||  | || ||    || ||||||||||  || |||| |||||  | |||||| || |||
Db        259 GGTTTATGCCAAATTGTTGGCTGCGATCACCAGCTGGCAAGCACCGTCAAGGAAGATAAC 318

Qy        649 TGTGGAGTCTGTGCCGGCGATGGCTCCACCTGCAGGCTTGTACGGGGACAATCAAAGTCA 708
              ||||| |||||   ||| ||||| ||||||||| |||| || || || || |  || ||
Db        319 TGTGGGGTCTGCAACGGAGATGGGTCCACCTGCCGGCTGGTCCGAGGGCAGTATAAATCC 378

Qy        709 CACGTTTCTCCTGAAAAAAGAGAAGAAAATGTAATTGCTGTTCCTTTGGGAAGTCGAAGT 768
              ||  | ||     | |||   || || | |||  ||||  |||| |  |||||| ||  |
Db        379 CAGCTCTCCGGCAACAAATCGGATGATACTGTGGTTGCAATTCCCTATGGAAGTAGACAT 438

Qy        769 GTGAGAATTACAGTGAAAGGACCTGCCCACCTCTTTATTGAATCAAAAACACTTCAAGGA 828
               |  |  ||    | ||||| ||||  ||| | | | | ||| | ||||| || || ||
Db        439 ATTCGCCTTGTCTTAAAAGGTCCTGATCACTTATATCTGGAAACCAAAACCCTCCAGGGG 498

Qy        829 AGCAAAGGAGAACACAGCTTTAACAGCCCCGGCGTCTTTGTCGTAGAAAACACAACAGTG 888
              |  ||||| ||| ||||  | | |  | | ||    ||  | || || ||  | |  |||
Db        499 ACTAAAGGTGAAAACAGTCTCAGCTCCACAGGAACTTTCCTTGTGGACAATTCTAGTGTG 558

Qy        889 GAATTTCAGAGGGGCTCCGAGAGGCAAACTTTTAAGATTCCAGGACCTCTGATGGCTGAT 948
              || || ||||      | || |   | |   | |  ||  | ||||| || |  || |||
Db        559 GACTTCCAGAAATTTCCAGACAAAGAGATACTGAGAATGGCTGGACCACTCACAGCAGAT 618

Qy        949 TTCATCTTCAAGACCAGGTACACTGCAGCCAAAGACAGCGTGGTTCAGTTCTTCTTTTAC 1008
              |||||  ||||||   |  || | |   ||   |||||    || |||||| |||| ||
Db        619 TTCATTGTCAAGATTCGTAACTCGGGCTCCGCTGACAGTACAGTCCAGTTCATCTTCTAT 678

Qy       1009 CAGCCCATCAGTCATCAGTGGAGACAAACTGACTTCTTTCCCTGCACTGTGACGTGTGGA 1068
              || |||||||  || |  |||||  | || || || ||||| ||| | |  || ||||||
Db        679 CAACCCATCATCCACCGATGGAGGGAGACGGATTTTTTTCCTTGCTCAGCAACCTGTGGA 738

Qy       1069 GGAGGTTATCAGCTCAATTCTGCTGAATGTGTGGATATCCGCTTGAAGAGGGTAGTTCCT 1128
              |||||||||||||| |  || ||||| ||    ||| |  |    ||  | || || |||
Db        739 GGAGGTTATCAGCTGACATCGGCTGAGTGCTACGATCTGAGGAGCAACCGTGTGGTTGCT 798

Qy       1129 GACCATTATTGTCACTACTACCCTGAAAATGTAAAACCAAAACCAAAACTGAAGGAATGC 1188
              ||||| || |||||||| ||||| || ||  | ||||| ||||| || ||  |||| |||
Db        799 GACCAATACTGTCACTATTACCCAGAGAACATCAAACCCAAACCCAAGCTTCAGGAGTGC 858

Qy       1189 AGCATGGATCCCTGCCCATCAAGTGATGGATTTAAAGAGATAATGCCCTATGACCACTTC 1248
              | | ||||||| || ||| | |
Db        859 AACTTGGATCCTTGTCCAGCCA-------------------------------------- 880

Qy       1249 CAACCTCTTCCTCGCTGGGAACATAATCCTTGGACTGCATGTTCCGTGTCCTGTGGAGGA 1308
                           | |||||    |  || |||||  | || |||   || ||||| ||
Db        881 -------------GGTGGGAGGCCACCCCATGGACCCCGTGCTCCTCCTCGTGTGGGGGC 927

Qy       1309 GGGATTCAGAGACGGAGCTTTGTGTGTGTAGAGGAATCCATGCATGGAGAGATATTGCAG 1368
              |  || ||||| |||    ||   ||||| |||||   ||| || ||  |  |
Db        928 GCCATCCAGAGCCGGGCAGTTTCCTGTGTGGAGGAGGACATCCAGGGGCATGTCACTTCA 987

Qy       1369 GTGGAAGAATGGAAGTGCATGTACGCACCCAAACCCAAGGTTATGCAAACTTGTAATCTG 1428
              |||||||| ||||| ||||||||| | || ||        |   |||  | || ||  |
Db        988 GTGGAAGAGTGGAAATGCATGTACACCCCTAAGATGCCCATCGCGCAGCCCTGCAACATT 1047

Qy       1429 TTTGATTGCCCCAAGTGGATTGCCATGGAGTGGTCTCAGTGCACAGTGACTTGTGGCCGA 1488
              ||||| ||||| || ||| | ||   ||||||||||| |||||||||||| |||||||
Db       1048 TTTGACTGCCCTAAATGGCTGGCACAGGAGTGGTCTCCGTGCACAGTGACATGTGGCCAG 1107

Qy       1489 GGGTTACGGTACCGGGTTGTTCTGTGTATTAACCACCGCGGAGAGCATGTTGGGGGCTGC 1548
              ||  |  | ||||| || || ||:|| ||  |||| || |||  |||    || |||||
Db       1108 GGCCTCAGATACCGTGTGGTCCTSTGCATCGACCATCGAGGAATGCACACAGGAGGCTGT 1167

Qy       1549 AATCCACAACTGAAGTTACACATCAAAGAAGAATGTGTCATTCCCATCCCGTGTTATAAA 1608
              |  ||| ||   |||   ||||| ||||| |||||  || | ||||  || || ||||||
Db       1168 AGCCCAAAAACAAAGCCCCACATAAAAGAGGAATGCATCGTACCCACTCCCTGCTATAAA 1227

Qy       1609 CCAAAAGAAAAAAGTCCAGTGCAAGCAAAATTGCCTTGGCTGAAACAAGCACAAGAACTA 1668
              || ||||| |||  ||||||  | || || ||||| ||| | |||||||| ||||| |||
Db       1228 CCCAAAGAGAAACTTCCAGTCGAGGCCAAGTTGCCATGGTTCAAACAAGCTCAAGAGCTA 1287

Qy       1669 GAAGAGACCAGAATAGCAACAGAAGAACCAACGTTCATTCCAGAACCCTGGTCAGCCTGC 1728
              |||||          |   |||| || ||  |||||||::||:| |||||||| ||||||
Db       1288 GAAGAAGGAGCTGCTGTGTCAGAGGAGCCCTCGTTCATYYCARAGGCCTGGTCGGCCTGC 1347

Qy       1729 AGTACCACGTGTGGGCCAGGTGTGCAGGTCCGCGAGGTGAAGTGCCGTGTGCTCCTCACA 1788
              |    ||| |||||    ||    ||:|| ||    || | |||||  ||||||||  |
Db       1348 ACAGTCACCTGTGGTGTGGGGACCCARGTGCGAATAGTCAGGTGCCAGGTGCTCCTGTCT 1407

Qy       1789 TTCACGCAGACTGAGACTGAGCTGCCCGAGGAAGAGTGTGAAGGCCCCAAGCTGCCCACC 1848
              ||| | ||| | | | |||| |||||    || |||||||||||||||||||   |  ||
Db       1408 TTCTCTCAGTCCGTGGCTGACCTGCCTATTGACGAGTGTGAAGGCCCCAAGCCAGCATCC 1467

Qy       1849 ---------------GAACGGCCCTGCCTCCTGGAAGCATGTGATGAGAGCCCGGCCTCC 1893
                             | | | || |||  |  ||||     |||    | ||| | |
Db       1468 CAGCGTGCCTGTTATGCAGGCCCATGCAGCGGGGAAATTCCTGAGTTCAACCCAGACGAG 1527

Qy       1894 CGAGAGCTACACATCCCTCTCCCTGAGGACAGTGACACGACTTACGACTGGGAGTACCCT 1953
                |||    | | |   |  ||   ||||    |||  |   || |||||||||||
Db       1528 ACAGATGGGCTCTTTGGTGGCCTGCAGGATTTCGACGAGCTGTATGACTGGGAGTATGAG 1587

Qy       1954 GGGTTCACCCCTTGCACAGCAACATGCCTTGGAGGCCATCAAGAAGCCATAGCAGTGTGC 2013
              |||||||||   ||| | |   | ||    |||||    || || ||  | |     |||
Db       1588 GGGTTCACCAAGTGCTCCGAGTCCTGTGGAGGAGGTGTCCAGGAGGCTGTGGTGAGCTGC 1647

Qy       2014 TTACATATCCAGACCCAGCACACAGTCAATGACAGCTTGTGTGATATGGTCCACCGTCCT 2073
              ||  | |  ||||| | | |  | |   | || | | |||| |  |    || ||| ||
Db       1648 TTGAACAAACAGACTCGGGAGCCTGCTGAGGAGAACCTGTGCGTGACCAGCCGCCGGCCC 1707

Qy       2074 CCAGCCATGAGCCAGGCCTGTAACACAGAGCCCTGTCCCCCCAGGTGGCATGTGGGCTCT 2133
              |||    |     || |||| ||    || ||||| ||  | |||||| |  | |||
Db       1708 CCACAGCTCCTGAAGTCCTGCAATTTGGATCCCTGCCCAGCAAGGTGGGAAATTGGCAAG 1767
```

```
Qy       2134 [illegible] 2193

Db       1768 [illegible] 1827

Qy       2194 CACC [illegible] 2235

Db       1828 [illegible] 1887

Qy       2236 GAAAA [illegible] 2295

Db       1888 [illegible] 1947

Qy       2296 GAAGAA [illegible] 2355

Db       1948 [illegible] 2007

Qy       2356 [illegible] 2415

Db       2008 [illegible] 2067

Qy       2416 [illegible] 2475

Db       2068 [illegible] 2127

Qy       2476 [illegible] 2535

Db       2128 [illegible] 2187

Qy       2536 [illegible] 2595

Db       2188 [illegible] 2247

Qy       2596 [illegible] 2655

Db       2248 [illegible] 2295

Qy       2656 GAGATGAAGACAAAACTTGGTGAGCAGGGTCCGCAGATCCTCAGTGTCCAGAGAGTCTAC 2715

Db       2296 AGGCCCGGGCGGCCATCCACGAAGCACAGCCCGCACATCGCGGCCGCCAGGAAGGTCTAC 2355

Qy       2716 ATTCAGACAAGGGAAGAGAAGGGTATTAACCTGACCATTGGTAGCAGAGCCTATTTGCTG 2775

Db       2356 ATACAGACTCGGCAGGCAGAGGAAGCTGCACTTCGTGGTGGGGGGCTTCGCCTACCTGCTC 2415

Qy       2776 [illegible] 2835

Db       2416 [illegible] 2475

Qy       2836 [illegible] 2895

Db       2476 [illegible] 2535

Qy       2896 [illegible] 2955

Db       2536 [illegible] 2595

Qy       2956 [illegible] 3015

Db       2596 [illegible] 2655

Qy       3016 CC 3017
              |
Db       2656 CC 2657
```

Search completed: March 22, 2003, 13:27:50
Job time : 340 secs

```
                      GenCore version 5.1.4_p5_4578
                 Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on:          March 22, 2003, 10:52:29 ; Search time 60 Seconds
                                            (without alignments)
                                            3755.447 Million cell updates/sec

Title:           US-09-784-358-2
Perfect score:   9346
Sequence:        1 MASWISPWWVLIGMVFMHSP..........LNDCSLDRYKQRCYQSCQEG 1691

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters:      908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries

Database :        A_Geneseq_101002:*
                  1:  [illegible]
                  2:  [illegible]
                  3:  [illegible]
                  4:  [illegible]
                  5:  [illegible]
                  6:  [illegible]
                  7:  [illegible]AA1986.DAT:*
                  8:  [illegible]
                  9:  [illegible]AA1988.DAT:*
                  10: [illegible]
                  11: [illegible]
                  12: [illegible]
                  13: [illegible]
                  14: [illegible]
                  15: [illegible]
                  16: [illegible]
                  17: [illegible]
                  18: [illegible]
                  19: [illegible]
                  20: [illegible]
                  21: [illegible]
                  22: [illegible]
                  23: [illegible]

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                              SUMMARIES
                      %
Result               Query
  No.     Score    Match Length DB   ID                             Description

    1      9346    100.0   1691  22  AAE07863                       Novel human protei
    2      9339     99.9   1691  22  AAE10609                       Human novel KIAA12
    3      8882     95.0   1617  22  AAE07870                       Novel human protei
    4      7066     75.6   1283  22  AAE10610                       Human novel KIAA12
    5      7060     75.5   1281  22  AAE10608                       Human novel KIAA12
    6      5086     54.4    926  22  AAB83379                       NOV14 protein sequ
    7      4693     50.2    886  21  AAB40500                       Human ORFX ORF264
    8      4689     50.2    845  22  AAE07868                       Novel human prote.
    9      4225     45.2    771  22  AAE07869                       Novel human prote.
   10      3956     42.3    724  22  AAE07866                       Novel human protei
```

```
   11      3861     41.3   1745  22  AAB90[illegible]               Human secreted pro
   12      3861     41.3   1745  23  ABG65473                       Human albumin fusi
   13    3860.5     41.3   1766  22  ABB10246                       Human cDNA SEQ ID
   14      3492     37.4    650  22  AAE07867                       Novel human protei
   15      2891     30.9    874  21  AAB01431                       Human TANGO 224 (I
   16    2351.5     25.2    643  21  AAY87347                       Human signal pepti
   17      2344     25.1    446  22  AAE07864                       Novel human protei
   18      2058     22.0    525  20  AAY41729                       Human PRO1071 prot
   19      2058     22.0    525  22  AAG29241                       Human PRO polypept
   20      2058     22.0    525  23  ABB95494                       Human angiogenesis
   21      2058     22.0    525  23  ABB84888                       Human PRO1071 prot
   22      2058     22.0    525  23  AAU83668                       Human PRO protein,
   23      2051     21.9    525  21  AAB44285                       Human PRO1071 (UNQ
   24      2048     21.9    525  22  AAB72290                       Human ADAMTS-E1 am
   25      1880     20.1    372  22  AAE07865                       Novel human protei
   26      1560     16.7    287  22  ABB11315                       Human HSPP 124 pro
   27    1536.5     16.4    428  21  AAB41561                       Human ORFX ORF1325
   28      1483     15.9    417  22  AAG78679                       Human thrombotic p
   29      1362     14.6    480  21  AAB01425                       Human TANGO 224.
   30    1183.5     12.7   2150  21  AAY53898                       Amino acid sequenc
   31    1183.5     12.7   2165  22  AAB90617                       Human secreted pro
   32      1044     11.2   1907  23  AAU77133                       Human protease #12
   33      1044     11.2   1935  23  AAU72896                       Human metalloprote
   34      1043     11.2   1916  23  AAE19173                       Human protease, PR
   35      1043     11.2   1934  22  AAB72301                       Human ADAMTS-9 alt
   36      1028     11.0   1882  22  AAB72289                       Human ADAMTS-9 ami
   37       972     10.4   1629  23  ABG30703                       Human aggrecanase
   38       972     10.4   1629  23  AAU14448                       Human ADAMTS-S1 pr
   39     957.5     10.2   1602  23  ABG30702                       Human aggrecanase
   40     938.5     10.0    262  22  AAB90587                       Human secreted pro
   41     938.5     10.0    262  23  ABG65471                       Human albumin fusi
   42     923.5      9.9   1535  23  AAU72897                       Human metalloprote
   43     920.5      9.8   1281  22  AAB11779                       Human ORFX ORF1143
   44   [illegible] [illegible] [illegible] 22 AAB80949             Human metalloprote
   45       892      9.5   1686  22  AAB71[illegible]               Human ADAM type me

                              ALIGNMENTS

RESULT 1
AAE07863
ID   AAE07863 standard; Protein; 1691 AA.
XX
AC   AAE07863;
XX
DT   21-NOV-2001  (first entry)
XX
DE   Novel human protein (NHP) #1.
XX
KW   Novel human protein; NHP; thrombospondin; gene therapy; cancer;
KW   cytostatic; antisense therapy; angiogenesis; biological disorder.
XX
OS   Homo sapiens.
XX
PN   WO200161011-A2.
XX
PD   23-AUG-2001.
XX
PF   15-FEB-2001; 2001WO-US05290.
XX
PR   17-FEB-2000; 2000US-0183282.
XX
PA   (LEXI-) LEXICON GENETICS INC.
XX
PI   Donoho G,  Scoville J,  Turner CA,  Friedrich G,  Zambrowicz B;
PI   Sands AT;
XX
DR   WPI; 2001-514776/56.
DR   N-PSDB; AAD14364.
XX
PT   An isolated nucleic acid encoding a thrombospondin useful as a
PT   hybridization probe and gene therapy treatments of cancer
```

```
XX
PS   Claim 2; Page 31-34; 56pp; English.
XX
CC   The invention relates to novel human polynucleotides encoding proteins
CC   that share sequence similarity with animal proteins having thrombospondin
CC   repeats. NHPs are expressed in human cell lines, pituitary, lymph nodes,
CC   prostrate, testis, adrenal gland, uterus, foetal kidney, foetal lung and
CC   gene trapped human cells. NHPs are thrombospondins useful for treating
CC   biological disorders including angiogenesis, cancer and development and
CC   also in pharmacogenomic applications. NHPs are useful as a hybridisation
CC   probe for screening libraries, assessing gene expression patterns and
CC   also in gene therapy. Proteins having thrombospondin repeats act as
CC   receptors, secreted extracellular matrix proteins and proteases. The
CC   present sequence is novel human protein (NHP).
XX
SQ   Sequence   1691 AA;

  Query Match             100.0%;  Score 9346;  DB 22;  Length 1691;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1691;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         1 MASWTSPWWVLIGMVFMHSPLPQTTAFKSPGAYFLPEFALSPQGSFLEDTTGEQFLTYRY 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 MASWTSPWWVLIGMVFMHSPLPQTTAFKSPGAYFLPEFALSPQGSFLEDTTGEQFLTYRY 60

Qy        61 DDQTSRNTRSDEDKDGNWDAWGDWSDCSRTCGGGASYSLRRCLTGRNCEGQNIRYKTCSN 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 DDQTSRNTRSDEDKDGNWDAWGDWSDCSRTCGGGASYSLRRCLTGRNCEGQNIRYKTCSN 120

Qy       121 HDCPPDAEDFRAQQCSAYNDVQYQGHYYEWLPRYNDPAAPCALKCHAQGQNLVVELAPKV 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 HDCPPDAEDFRAQQCSAYNDVQYQGHYYEWLPRYNDPAAPCALKCHAQGQNLVVELAPKV 180

Qy       181 LDGTRCNTDSLDMCISGICQAVGCDRQLGSNAKEDNCGVCAGDGSTCRLVRGQSKSHVSP 240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       181 LDGTRCNTDSLDMCISGICQAVGCDRQLGSNAKEDNCGVCAGDGSTCRLVRGQSKSHVSP 240

Qy       241 EKREENVIAVPLGSRSVRITVKGPAHLFIESKTLQGSKGEHSFNSPGVFVVENTTVEFQR 300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       241 EKREENVIAVPLGSRSVRITVKGPAHLFIESKTLQGSKGEHSFNSPGVFVVENTTVEFQR 300

Qy       301 GSERQTFKIPGPLMADFIFKTRYTAAKDSVVQFFFYQPISHQWRQTDFFPCTVTCGGGYQ 360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       301 GSERQTFKIPGPLMADFIFKTRYTAAKDSVVQFFFYQPISHQWRQTDFFPCTVTCGGGYQ 360

Qy       361 LNSAECVDIRLKRVVPDHYCHYYPENVKPKPKLKECSMDPCPSSDGFKEIMPYDHFQPLP 420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       361 LNSAECVDIRLKRVVPDHYCHYYPENVKPKPKLKECSMDPCPSSDGFKEIMPYDHFQPLP 420

Qy       421 RWEHNPWTACSVSCGGGIQRRSFVCVEESMHGEILQVEEWKCMYAPKPKVMQTCNLFDCP 480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       421 RWEHNPWTACSVSCGGGIQRRSFVCVEESMHGEILQVEEWKCMYAPKPKVMQTCNLFDCP 480

Qy       481 KWIAMEWSQCTVTCGRGLRYRVVLCINHRGEHVGGCNPQLKLHIKEECVIPIPCYKPKEK 540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       481 KWIAMEWSQCTVTCGRGLRYRVVLCINHRGEHVGGCNPQLKLHIKEECVIPIPCYKPKEK 540

Qy       541 SPVEAKLPWLKQAQELEETRIATEEPTFIPEPWSACSTTCGPGVQVREVKCRVLLTFTQT 600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       541 SPVEAKLPWLKQAQELEETRIATEEPTFIPEPWSACSTTCGPGVQVREVKCRVLLTFTQT 600

Qy       601 ETELPEEECEGPKLPTERPCLLEACDESPASRELDIPLPEDSETTYDWEYAGFTPCTATC 660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       601 ETELPEEECEGPKLPTERPCLLEACDESPASRELDIPLPEDSETTYDWEYAGFTPCTATC 660

Qy       661 LGGHQEATAVCLHIQTQQTVNDSLCDMVHRPPAMSQACNTEPCPPRWHVGSWGPCSATCG 720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       661 LGGHQEATAVCLHIQTQQTVNDSLCDMVHRPPAMSQACNTEPCPPRWHVGSWGPCSATCG 720

Qy       721 VGIQTRDVYCLHPGETPAPPEECRDEKPHALQACNQFDCPPGWHIEEWQQCSRTCGGGTQ 780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       721 VGIQTRDVYCLHPGETPAPPEECRDEKPHALQACNQFDCPPGWHIEEWQQCSRTCGGGTQ 780

Qy       781 NRRVTCRQLLTDGSFLNLSDELCQGPKASSHKSCARTDCPPHLAVGDWSKCSVSCGVGIQ 840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       781 NRRVTCRQLLTDGSFLNLSDELCQGPKASSHKSCARTDCPPHLAVGDWSKCSVSCGVGIQ 840

Qy       841 RPFQVCQRLAAKCRRIPLSEMMCRDLPGFPLVRSCQMPECSKIESEMKIKLCEQCPQILS 900
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       841 RPFQVCQRLAAKCRRIPLSEMMCRDLPGFPLVRSCQMPECSKIESEMKIKLGEQGPQILS 900

Qy       901 VQPVYIQTEEEKKINLLQGSEAYLLENTSVIIKCTVEEFQKCLQWEKLGKCLQNSKRLG 960
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       901 VQPVYIQTEEEKKINLLQGSEAYLLENTSVIIKCTVEEFQKCLQWEKLGKCLQNSKRLG 960

Qy       961 ITKSGSLKIHGLAAFDIGVYRCIACSAQETVVLKLLIGTDNRLIARPALREPMREYPGMDH 1020
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       961 ITKSGSLKIHGLAAFDIGVYRCIAGSAQETVVLKLLIGTDNRLIARPALREPMREYPGMDH 1020

Qy      1021 SEANSLGVTWHKMRQMWNNKNDLYLLEDHIISNQPFLRALLIGHCSNSAGSTNSWELKNEQF 1080
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1021 SEANSLGVTWHKMRQMWNNKNDLYLLEDHIISNQPFLRALLIGHCSNSAGSTNSWELKNEQF 1080

Qy      1081 EAAVKQGAYSMDTAQFDELIRNMSQLMETGEVSDDLASQLIYQLVAELAKAQPTHMQWRG 1140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1081 EAAVKQGAYSMDTAQFDELIRNMSQLMETGEVSDDLASQLIYQLVAELAKAQPTHMQWRG 1140

Qy      1141 IQEETPPAAQLRGETGSVSQSSHAKNSGKLTFKPKGPVLMRQSQPPSISFNKTINSRIGN 1200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1141 IQEETPPAAQLRGETGSVSQSSHAKNSGKLTFKPKGPVLMRQSQPPSISFNKTINSRIGN 1200

Qy      1201 TVYITKRTEVINILCDLITPSEATYTWTKDGTLLQPSVKIILDGTGKIQIQNPTRKEQGI 1260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1201 TVYITKRTEVINILCDLITPSEATYTWTKDGTLLQPSVKIILDGTGKIQIQNPTRKEQGI 1260

Qy      1261 YECSVANHLGSDVESSSVLYAEAPVILSVERNITKPEHNHLSVVVGGIVEAALGANVTIR 1320
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1261 YECSVANHLGSDVESSSVLYAEAPVILSVERNITKPEHNHLSVVVGGIVEAALGANVTIR 1320

Qy      1321 CPVKGVPQPNITWLKRGGSLSGNVSLLFNGSLLLQNVSLENEGTYVCIATNALGKAVAIS 1380
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1321 CPVKGVPQPNITWLKRGGSLSGNVSLLFNGSLLLQNVSLENEGTYVCIATNALGKAVAIS 1380

Qy      1381 VLHLLERRWPESRIVFLQGHKKYILQATNTRTNSNDPTGEPPPQEPFWEPGNWSHCSATC 1440
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1381 VLHLLERRWPESRIVFLQGHKKYILQATNTRTNSNDPTGEPPPQEPFWEPGNWSHCSATC 1440

Qy      1441 CHLCARIQRPQCVMANGQEVSEALCDHLQKPLAGFEPCNIRDCPARWFTSVWSQCSVSCG 1500
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1441 GHLGARIQRPQCVMANGQEVSEALCDHLQKPLAGFEPCNIRDCPARWFTSVWSQCSVSCG 1500

Qy      1501 EGYHSRQVTCKRTKANGTVQVVSPRACAPKDRPLGRKPCFGHPCVQWEPGNRCPGRCMGR 1560
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1501 EGYHSRQVICKRTKANGTVQVVSPRACAPKDRPLGRKPCFGHPCVQWEPGNRCPGRCMGR 1560

Qy      1561 AVRMQQRHTACQHNSSDSNCDDRKRPTLRRNCTSGACDVCWHTGPWKPCTAACGRGFQSR 1620
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1561 AVRMQQRHTACQHNSSDSNCDDRKRPTLRRNCTSGACDVCWHTGPWKPCTAACGRGFQSR 1620

Qy      1621 KVDCIHTRSCKPVAKRHCVQKKKPISWRHCLGPSCDRDCTDITHYCMFVKHLNLCSLDRY 1680
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1621 KVDCIHTRSCKPVAKRHCVQKKKPISWRHCLGPSCDRDCTDITHYCMFVKHLNLCSLDRY 1680

Qy      1681 KQRCCQSCQEG 1691
             |||||||||||
Db      1681 KQRCCQSCQEG 1691


RESULT 2
AAE10609
ID   AAE10609 standard; Protein; 1691 AA.
XX
AC   AAE10609;
XX
DT   10-DEC-2001  (first entry)
XX
```

```
DE   Human novel KIAA1233 like protein, NoV 2a.
XX
KW   Human; NoV X protein; KIAA1233 like protein; STE20 like protein; tumour;
KW   trypsin inhibitor like protein; gene therapy; haematopoietic; illness;
KW   immunological disorder; neurodegenerative disorder; Alzheimer's disease;
KW   Parkinson's disease; immunomodulatory; pharmacogenomic; haemostatic;
KW   human immunodeficiency virus; HIV; fertility disorder; neuroprotective;
KW   cytostatic; nootropic; anti infertility; cancer; chromosome 15;
KW   NoV 2a protein.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Active site     1..26
FT                   /label= Signal peptide
FT   Protein         27..1691
FT                   /note= "Human mature novel KIAA1233 like protein,
FT                   NoV 2a"
XX
PN   WO200162928-A2.
XX
PD   30-AUG-2001.
XX
PF   26-FEB-2001; 2001WO-US06151.
XX
PR   25-FEB-2000; 2000US-0184951.
PR   28-FEB-2000; 2000US-0185118.
PR   01-MAR-2000; 2000US-0185962.
PR   18-APR-2000; 2000US-0197723.
PR   27-APR-2000; 2000US-0200667.
PR   23-FEB-2001; 2001US-0790568.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Vernet CAM,  Fernandes E,  Shimkets RA,  Macdougall J,  Spaderna SK;
XX
DR   WPI; 2001-582051/65.
DR   N-PSDB; AAD17757.
XX
PT   New isolated KIAA1233 like, STE20 like, or trypsin inhibitor-like
PT   polypeptide for diagnosing and treating pathological disorders, such as
PT   Parkinson's disease and for use in pharmacogenomics
XX
PS   Claim 1; Page 22-23; 180pp; English.
XX
CC   The invention relates to novel human polypeptides referred as NoV X
CC   and their corresponding nucleic acid sequences. NoV X collectively
CC   include NoV 1, NoV 2a and NoV 2b which are novel KIAA1233 like
CC   polypeptides, NoV 3a, NoV 3b, NoV 3c and NoV 3d which are novel
CC   STE20 like polypeptides and NoV 4a, NoV 4b, NoV 4c, NoV 4d and NoV 4e
CC   which are novel trypsin inhibitor like polypeptides. NoV X is used to
CC   identify a potential therapeutic agent that can modulate its activity
CC   and can be used for treating a pathology related to aberrant expression
CC   or aberrant physiological interactions of NoV X. NoV X or its DNA is
CC   used to determine the presence or predisposition to a disease associated
CC   with altered levels of NoV X. NoV X, its DNA and its antibody are used
CC   to treat or prevent a pathology associated with NoV X. The pathological
CC   states that can be treated or prevented are haematopoietic, cancer,
CC   immunological, tumour, neurodegenerative (e.g. Alzheimer's and
CC   Parkinson's disease), human immunodeficiency virus (HIV) illness and
CC   fertility disorders. NoV X and its DNA are used in pharmacogenomics for
CC   predictive medicine. NoV X DNA is used in gene therapy. The present
CC   sequence is human novel KIAA1233 like protein, NoV 2a. NoV 2a gene is
CC   located at chromosome 15.
XX
SQ   Sequence   1691 AA;

  Query Match             99.4%;  Score 9339;  DB 22;  Length 1691;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1689;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MASWISPWWVLIGMVFMHSPLPQIIAEKSPGAYFLPEFALSPQGSFLEDTTGEQFLTYRY 60
              |||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MASWISPWWVLIGMVFMHSPLPQIIAEKSPGAYFLPEFALSPQGSFLEDTTGEQFLTYRY 60

Qy         61 DDQISRNTRSDEDKDGNWDAWGDWSDCSRTCGGGASYSLRRCLTGRNCEGQNIRYKTCSN 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DDQISRNTRSDEDKDGNWDAWGDWSDCSRTCGGGASYSLRRCLTGRNCEGQNIRYKTCSN 120

Qy        121 HDCPPDAEDFRAQQCSAYNDVQYQGHYYEWLPRYNDPAAPCALKCHAQGQNLVVELAPKV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 HDCPPDAEDFRAQQCSAYNDVQYQGHYYEWLPRYNDPAAPCALKCHAQGQNLVVELAPKV 180

Qy        181 LDGTRCNTDSLDMCISGICQAVGCDRQLGSNAKEDNCGVCAGDGSTCRLVRGQSKSHVSP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 LDGTRCNTDSLDMCISGICQAVGCDRQLGSNAKEDNCGVCAGDGSTCRLVRGQSKSHVSP 240

Qy        241 EKREENVIAVPLGSRSVRITVKGPAHLFIESKTLQGSKGEHSFNSPGVFVVENTTVEFQR 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 EKREENVIAVPLGSRSVRITVKGPAHLFIESKTLQGSKGEHSFNSPGVFVVENTTVEFQR 300

Qy        301 GSERQTFKIPGPLMADFIFKTRYTAAKDSVVQFFFYQPISHQWRQTDFFPCTVTCGGGYQ 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GSERQTFKIPGPLMADFIFKTRYTAAKDSVVQFFFYQPISHQWRQTDFFPCTVTCGGGYQ 360

Qy        361 LNSAECVDIRLKRVVPDHYCHYYPENVKPKPKLKECSMDPCPSSDGFKEIMPYDHFQPLP 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LNSAECVDIRLKRVVPDHYCHYYPENVKPKPKLKECSMDPCPSSDGFKEIMPYDHFQPLP 420

Qy        421 PWEHNPWTACSVSCGGGIQRRSFVCVEESMHGEILQVEEWKCMYAPKPKVMQTCNLFDCP 480
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 RWEHNPWTACSVSCGGGIQRRSFVCVEESMHGEILQVEEWKCMYAPKPKVMQTCNLFDCP 480

Qy        481 KWIAMEWSQCTVTCGRGLRYRVVLCINHRGEHVGGCNPQLKLHIKEECVIPIPCYKPKEK 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 KWIAMEWSQCTVTCGRGLRYRVVLCINHRGEHVGGCNPQLKLHIKEECVIPIPCYKPKEK 540

Qy        541 SPVEAKLPWLKQAQELEETRIATEEPTFIPEPWSACSTTCGPGVQVREVKCRVLLTFTQT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 SPVEAKLPWLKQAQELEETRIATEEPTFIPEPWSACSTTCGPGVQVREVKCRVLLTFTQT 600

Qy        601 ETELPEEECEGPKLPTERPCLLEACDESPASRELDIPLPEDSETTYDWEYAGFTPCTATC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ETELPEEECEGPKLPTERPCLLEACDESPASRELDIPLPEDSETTYDWEYAGFTPCTATC 660

Qy        661 LGGHQEAIAVCLHIQTQQTVNDSLCDMVHRPPAMSQACNTEPCPPRWHVGSWGPCSATCG 720
              :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 VGGHQEAIAVCLHIQTQQTVNDSLCDMVHRPPAMSQACNTEPCPPRWHVGSWGPCSATCG 720

Qy        721 VGIQTRDVYCLHPGETPAPPEECRDEKPHALQACNQFDCPPGWHIEEWQQCSRTCGGGTQ 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 VGIQTRDVYCLHPGETPAPPEECRDEKPHALQACNQFDCPPGWHIEEWQQCSRTCGGGTQ 780

Qy        781 NRRVTCRQLLTDGSFLNLSDELCQGPKASSHKSCARTDCPPHLAVGDWSKCSVSCGVGIQ 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 NRRVTCRQLLTDGSFLNLSDELCQGPKASSHKSCARTDCPPHLAVGDWSKCSVSCGVGIQ 840

Qy        841 RRKQVCQRLAAKGRRIPLSEMMCRDLPGFPLVRSCQMPECSKIKSEMKTKLGEQGPQILS 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 RRKQVCQRLAAKGRRIPLSEMMCRDLPGFPLVRSCQMPECSKIKSEMKTKLGEQGPQILS 900

Qy        901 VQRVYIQTREEKRINLTIGSRAYLLPNTSVIIKCPVRRFQKSLIQWEKDGRCLQNSKRLG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 VQRVYIQTREEKRINLTIGSRAYLLPNTSVIIKCPVRRFQKSLIQWEKDGRCLQNSKRLG 960

Qy        961 ITKSGSLKIHGLAAPDIGVYRCIAGSAQETVVLKLIGTDNRLIARPALREPMREYPGMDH 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 ITKSGSLKIHGLAAPDIGVYRCIAGSAQETVVLKLIGTDNRLIARPALREPMREYPGMDH 1020

Qy       1021 SEANSLGVTWHKMRQMWNNKNDLYLDDDHISNQPFLRALLGHCSNSAGSTNSWELKNKQF 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 SEANSLGVTWHKMRQMWNNKNDLYLDDDHISNQPFLRALLGHCSNSAGSTNSWELKNKQF 1080

Qy       1081 EAAVKQGAYSMDTAQFDELIRNMSQLMETGEVSDDLASQLIYQLVAELAKAQPTHMQWRG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 EAAVKQGAYSMDTAQFDELIRNMSQLMETGEVSDDLASQLIYQLVAELAKAQPTHMQWRG 1140
```

```
Qy    1141 IQEEIPPAAQLRGETGSVSQSSHAKNSGKLTFKFFGPVLMEQSLPPSISPNETINSRICN 1200
           ||||||||||||||||||||||||||||||||||||  ||||||||||||||||||||||
Db    1141 IQEEIPPAAQLRGETGSVSQSSHAKNSGKLTFKFFGPVLMEQSLPPSISPNETINSRICN 1200

Qy    1201 TVYIIKRTEVINILCDLITPSEATYTWTKDGTLLQPSVKIILDGTGKIQIQNPTRKEQGI 1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 TVYIIKRTEVINILCDLITPSEATYTWTKDGTLLQPSVKIILDGTGKIQIQNPTRKEQGI 1260

Qy    1261 YECSVANHLGSDVESSSVLYAEAPVILSVERNIIKPEHNHLSVVVGGIVEAALGANVTIR 1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 YECSVANHLGSDVESSSVLYAEAPVILSVERNIIKPEHNHLSVVVGGIVEAALGANVTIR 1320

Qy    1321 CPVKGVPQPNITWLKRGGSLSGNVSLLFNGSLLLQNVSLENEGIYVCIAINALGKAVATS 1380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1321 CPVKGVPQPNITWLKRGGSLSGNVSLLFNGSLLLQNVSLENEGIYVCIAINALGKAVATS 1380

Qy    1381 VLHLLERRWPESRIVFLQGHKKYILQATNTRTNSNDPTGEPPPQEPFWEPGNWSHCSATC 1440
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1381 VFHLLERRWPESRIVFLQGHKKYILQATNTRTNSNDPTGEPPPQEPFWEPGNWSHCSATC 1440

Qy    1441 GHLGARIQRPQCVMANGQEVSEALCDHLQKPLAGFEPCNIRDCPARWFTSVWSQCSVSCG 1500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1441 GHLGARIQRPQCVMANGQEVSEALCDHLQKPLAGFEPCNIRDCPARWFTSVWSQCSVSCG 1500

Qy    1501 EGYHSRQVTCKRTKANGIVQVVSPRACAPKDRPLGRKPCFGHPCVQWEPGNRCPGRCMGR 1560
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1501 EGYHSRQVTCKRTKANGIVQVVSPRACAPKDRPLGRKPCFGHPCVQWEPGNRCPGRCMGR 1560

Qy    1561 AVRMQQRHTACQHNSSDSNCDDRKRPTLRRNCTSGACDVCWHTGPWKPCIAACGRGFQSR 1620
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1561 AVRMQQRHTACQHNSSDSNCDDRKRPTLRRNCTSGACDVCWHTGPWKPCIAACGRGFQSR 1620

Qy    1621 KVDCIHTRSCKPVAKRHCVQKKKPISWRHCLGPSCDRDCTDTTHYCMFVKHLNLCSLDRY 1680
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1621 KVDCIHTRSCKPVAKRHCVQKKKPISWRHCLGPSCDRDCTDTTHYCMFVKHLNLCSLDRY 1680

Qy    1681 KQRCCQSCQEG 1691
           |||||||||||
Db    1681 KQRCCQSCQEG 1691


RESULT 3
AAE07870
ID   AAE07870 standard; Protein; 1617 AA.
XX
AC   AAE07870;
XX
DT   01-NOV-2001  (first entry)
XX
DE   Novel human protein (NHP) #8.
XX
KW   Novel human protein; NHP; thrombospondin; gene therapy; cancer;
KW   cytostatic; antisense therapy; angiogenesis; biological disorder
XX
OS   Homo sapiens.
XX
PN   WO200161011-A2.
XX
PD   23-AUG-2001.
XX
PF   15-FEB-2001; 2001WO-US05290.
XX
PR   17-FEB-2000; 2000US-0183282.
XX
PA   (LEXI-) LEXICON GENETICS INC.
XX
PI   Donoho G,  Scoville J,  Turner CA,  Friedrich G,  Zambrowicz B;
PI   Sands AT;
XX
DR   WPI; 2001-514776/56.
DR   N-PSDB; AAD14371.
XX
PT   An isolated nucleic acid encoding a thrombospondin useful as a
PT   hybridization probe and gene therapy treatments of cancer -
XX
PS   Claim 4; Page 50-55; 56pp; English.
XX
CC   The invention relates to novel human polynucleotides encoding proteins
CC   that share sequence similarity with animal proteins having thrombospondin
CC   repeats. NHPs are expressed in human cell lines, pituitary, lymph nodes,
CC   prostrate, testis, adrenal gland, uterus, foetal kidney, foetal lung and
CC   gene trapped human cells. NHPs are thrombospondins useful for treating
CC   biological disorders involving angiogenesis, cancer and development and
CC   also in pharmacogenomic applications. NHPs are useful as a hybridisation
CC   probe for screening libraries, assessing gene expression patterns and
CC   also in gene therapy. Proteins having thrombospondin repeats act as
CC   receptors, secreted extracellular matrix proteins and proteases. The
CC   present sequence is novel human protein (NHP).
XX
SQ   Sequence   1617 AA;

  Query Match             95.0%;  Score 8882;  DB 22;  Length 1617;
  Best Local Similarity   95.6%;  Pred. No. 0;
  Matches 1617;  Conservative    0;  Mismatches    0;  Indels   74;  Gaps    1;

Qy          1 MASWTSPWWVLIGMVFMHSPLPQTTAEKSPGAYFLPEFALSPQGSFLEDTTGEQFLTYRY 60
              ||||||||||||||||||||||||||||||
Db          1 MASWTSPWWVLIGMVFMHSPLPQTTAEKSP------------------------------ 30

Qy         61 DEQISFNIPSSELEDGNWPAWGSWSDCSRTCGGGASYSLRRCLTGRNCEGQNIRYKTCSN 120
                                                          ||||||||||||||||
Db         31 --------------------------------------------GRNCEGQNIRYKTCSN 46

Qy        121 HECPPDAEDFRAQQCSAYNDVQYQGHYYEWLPRYNDPAAPCALKCHAQGQNLVVELAPKV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         47 HECPPDAEDFRAQQCSAYNDVQYQGHYYEWLPRYNDPAAPCALKCHAQGQNLVVELAPKV 106

Qy        181 LDGTRCNTDSLDMCISGICQAVGCDRQLGSNAKEDNCCVCACDGSTCRLVRGQSKSHVSP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 LDGTRCNTDSLDMCISGICQAVGCDRQLGSNAKEDNCCVCACDGSTCRLVRGQSKSHVSP 166

Qy        241 EKREENVIAVPLGSRSVRITVKGPAHLFIESKTLQGSKGEHSFNSPGVFVVENTTVEFQR 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 EKREENVIAVPLGSRSVRITVKGPAHLFIESKTLQGSKGEHSFNSPGVFVVENTTVEFQR 226

Qy        301 GSERQTFKIPGPLMADFIFKTRYTAAKDSVVQFFFYQPISHQWRQTDFFPCTVTCGGGYQ 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 GSERQTFKIPGPLMADFIFKTRYTAAKDSVVQFFFYQPISHQWRQTDFFPCTVTCGGGYQ 286

Qy        361 LNSAECVDIRLKRVVPLHYCHYYPENVKPKPKLKECSMDPCPSSDGFKEIMPYDHFQPLP 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 LNSAECVDIRLKRVVPLHYCHYYPENVKPKPKLKECSMDPCPSSDGFKEIMPYDHFQPLP 346

Qy        421 RWEHNPWTACSVSCGGGIQRRSFVCVEESMHGEILQVEEWKCMYAPKPKVMQTCNLFDCP 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        347 RWEHNPWTACSVSCGGGIQRRSFVCVEESMHGEILQVEEWKCMYAPKPKVMQTCNLFDCP 406

Qy        481 KWIAMEWSQCTVTCGRGLRYRVVLCINHRGEHVGGCNPQTKLHIKEECVIPIPCYKPKEK 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        407 KWIAMEWSQCTVTCGRGLRYRVVLCINHRGEHVGGCNPQTKLHIKEECVIPIPCYKPKEK 466

Qy        541 SPVEAKLPWLKQAQELEETRIATEEPTFIPEPWSACSTTCGPGVQVREVKCRVLLTFTQT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        467 SPVEAKLPWLKQAQELEETRIATEEPTFIPEPWSACSTTCGPGVQVREVKCRVLLTFTQT 526

Qy        601 ETELPEEECEGPKLPTERPCLLEACDESPASRELDIPLPEDSETTYDWEYAGFTPCTATC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        527 ETELPEEECEGPKLPTERPCLLEACDESPASRELDIPLPEDSETTYDWEYAGFTPCTATC 586

Qy        661 LGGHQEATAVCLHIQTQQTVNDSLCDMVHRPPAMSQACNTEPCPPRWHVGSWGPCSATCG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        587 LGGHQEATAVCLHIQTQQTVNDSLCDMVHRPPAMSQACNTEPCPPRWHVGSWGPCSATCG 646

Qy        721 VGIQTRDVYCLHPGETPAPPEECRDEKPHALQACNQFDCPPGWHIEEWQQCSRTCGGGTQ 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        647 VGIQIRDVYCLHPGEIPAPEEDCRDEKPHALQACNQFDCTPGWHIEEWQQCSRICGGGIQ 706

Qy        781 NRRVICRQLLIDGSFLNLSDELCQGPKASSHKSCARTDCPTHLAVGDWSKCSVSCGVGIQ 840
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        707 NRRVICRQLLIDGSFLNLSDELCQGPKASSHKSCARTDCPTHLAVGDWSKCSVSCGVGIQ 766

Qy        841 RRKQVCQRLAAKGRRIPLSEMMCRDLPGFPLVRSCQMPECSKIKSEMKTKLGEQGPQILS 900
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        767 RRKQVCQRLAAKGRRIPLSEMMCRDLPGFPLVRSCQMPECSKIKSEMKTKLGEQGPQILS 826

Qy        901 VQRVYIQTREEKRINLTIGSRAYLLPNTSVIIKCPVRRFQKSLIQWEKDGRCLQNSKRLG 960
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        827 VQRVYIQTREEKRINLTIGSRAYLLPNTSVIIKCPVRRFQKSLIQWEKDGRCLQNSKRLG 886

Qy        961 ITKSGSLKIHGLAAPDIGVYRCIAGSAQETVVLKLIGTDNRLIARPALREPMREYPGMDH 1020
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        887 ITKSGSLKIHGLAAPDIGVYRCIAGSAQETVVLKLIGTDNRLIARPALREPMREYPGMDH 946

Qy       1021 SEANSLGVTWHKMRQMWNNKNDLYLDDDHISNQPFLRALLGHCSNSAGSTNSWELKNKQF 1080
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        947 SEANSLGVTWHKMRQMWNNKNDLYLDDDHISNQPFLRALLGHCSNSAGSTNSWELKNKQF 1006

Qy       1081 EAAVKQGAYSMDTAQFDELIRNMSQLMETGEVSDDLASQLIYQLVAELAKAQPTHMQWRG 1140
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1007 EAAVKQGAYSMDTAQFDELIRNMSQLMETGEVSDDLASQLIYQLVAELAKAQPTHMQWRG 1066

Qy       1141 IQEETPPAAQLRGETGSVSQSSHAKNSGKLTFKPKGPVLMRQSQPPSISFNKTINSRIGN 1200
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1067 IQEETPPAAQLRGETGSVSQSSHAKNSGKLTFKPKGPVLMRQSQPPSISFNKTINSRIGN 1126

Qy       1201 TVYITKRTEVINILCDLITPSEATYTWTKDGTLLQPSVKIILDGTGKISIQNPTRKEQGI 1260
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1127 TVYITKRTEVINILCDLITPSEATYTWTKDGTLLQPSVKIILDGTGKISIQNPTRKEQGI 1186

Qy       1261 YECSVANHLGSDVESSSVLYAEAPVILSVERNITKPEHNHLSVVVGGIVEAALGANVTIR 1320
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1187 YECSVANHLGSDVESSSVLYAEAPVILSVERNITKPEHNHLSVVVGGIVEAALGANVTIR 1246

Qy       1321 CPVKGVPQPNITWLKRGGSLSGNVSLLFNGSLLLQNVSLENEGTYVCIATNALGKAVATS 1380
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1247 CPVKGVPQPNITWLKRGGSLSGNVSLLFNGSLLLQNVSLENEGTYVCIATNALGKAVATS 1306

Qy       1381 VLHLLERRWPESRIVFLQGHKKYILQATNTRTNSNDPTGEPPPQEPFWEPGNWSHCSATC 1440
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1307 VLHLLERRWPESRIVFLQGHKKYILQATNTRTNSNDPTGEPPPQEPFWEPGNWSHCSATC 1366

Qy       1441 GHLGARIQRPQCVMANGQEVSEALCDHLQKPLAGFEPCNIRDCPARWFTSVWSQCSVSCG 1500
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1367 GHLGARIQRPQCVMANGQEVSEALCDHLQKPLAGFEPCNIRDCPARWFTSVWSQCSVSCG 1426

Qy       1501 EGYHSRQVTCKRTKANGTVQVVSPRACAPKDRPLGRKPCFGHPCVQWEPGNRCPGRCMGR 1560
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1427 EGYHSRQVTCKRTKANGTVQVVSPRACAPKDRPLGRKPCFGHPCVQWEPGNRCPGRCMGR 1486

Qy       1561 AVRMQQRHTACQHNSSDSNCDDRKRPTLRRNCTSGACDVCWHTGPWKPCTAACGRGFQSR 1620
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1487 AVRMQQRHTACQHNSSDSNCDDRKRPTLRRNCTSGACDVCWHTGPWKPCTAACGRGFQSR 1546

Qy       1621 KVDCIHTRSCKPVAKRHCVQKKKPISWRHCLQPSCSRDCTDTTHYCMPVKHLNLCSLEPY 1680
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1547 KVDCIHTRSCKPVAKRHCVQKKKPISWRHCLQPSCSRDCTDTTHYCMPVKHLNLCSLEPY 1606

Qy       1681 KQRCCQSCQEG 1691
              |||||||||||
Db       1607 KQRCCQSCQEG 1617


RESULT 4
AAE10610
ID   AAE10610 standard; Protein; 1281 AA.
XX
AC   AAE10610;
XX
DT   13-DEC-2001  (first entry)
XX
DE   Human novel KIAA1233 like protein, NOV 2b.
XX
KW   Human; NOV X protein; KIAA1233 like protein; STE20 like protein; tumour;
KW   trypsin inhibitor-like protein; gene therapy; haematopoietic; illness;
KW   immunological disorder; neurodegenerative disorder; Alzheimer's disease;
KW   Parkinson's disease; immunomodulatory; pharmacogenomic; haemostatic;
KW   human immunodeficiency virus; HIV; fertility disorder; neuroprotective;
KW   cytostatic; nootropic; anti-infertility; cancer; chromosome 15;
KW   NOV 2b protein.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Active-site     1..26
FT                   /label= Signal peptide
FT   Protein         27..1281
FT                   /note= "Human mature novel KIAA1233 like protein,
FT                   NOV 2b"
XX
PN   WO200162928-A2.
XX
PD   30-AUG-2001.
XX
PF   26-FEB-2001; 2001WO-US06151.
XX
PR   25-FEB-2000; 2000US-0184971.
PR   28-FEB-2000; 2000US-0185548.
PR   01-MAR-2000; 2000US-0185967.
PR   18-APR-2000; 2000US-0197733.
PR   27-APR-2000; 2000US-0199947.
PR   23-FEB-2001; 2001US-0789390.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Vernet CAM,  Fernandes E,  Shimkets RA,  Macdougall J,  Spaderna SK;
XX
DR   WPI; 2001-582053/65.
DR   N-PSDB; AAD17758.
XX
PT   New isolated KIAA1233 like, STE20 like, or trypsin inhibitor like
PT   polypeptide for diagnosing and treating pathological disorders, such as
PT   Parkinson's disease and for use in pharmacogenomics
AA
PS   Claim 1; Page 37; 189pp; English.
XX
CC   The invention relates to novel human polypeptides referred as NOV X
CC   and their corresponding nucleic acid sequences. NOV X collectively
CC   include NOV 1, NOV 2a and NOV 2b which are novel KIAA1233 like
CC   polypeptides, NOV 3a, NOV 3b, NOV 3c and NOV 3d which are novel
CC   STE20 like polypeptides and NOV 4a, NOV 4b, NOV 4c, NOV 4d and NOV 4e
CC   which are novel trypsin inhibitor like polypeptides. NOV X is used to
CC   identify a potential therapeutic agent that can modulate its activity
CC   and can be used for treating a pathology related to aberrant expression
CC   or aberrant physiological interactions of NOV X. NOV X or its DNA is
CC   used to determine the presence or predisposition to a disease associated
CC   with altered levels of NOV X. NOV X, its DNA and its antibody are used
CC   to treat or prevent a pathology associated with NOV X. The pathological
CC   states that can be treated or prevented are haematopoietic, cancer,
CC   immunological, tumour, neurodegenerative (e.g. Alzheimer's and
CC   Parkinson's disease), human immunodeficiency virus (HIV) illness and
CC   fertility disorders. NOV X and its DNA are used in pharmacogenomics for
CC   predictive medicine. NOV X DNA is used in gene therapy. The present
CC   sequence is human novel KIAA1233 like protein, NOV 2b. NOV 2b gene is
CC   located at chromosome 15.
XX
SQ   Sequence   1281 AA;

  Query Match             75.6%;  Score 7066;  DB 22;  Length 1281;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1280;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy     411 MPYDHFQPLPRWEHNPWTACSVSCGGGIQRRSFVCVEESMHGEILQVEEWKCMYAPKPKV 470
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MPYDHFQPLPRWEHNPWTACSVSCGGGIQRRSFVCVEESMHGEILQVEEWKCMYAPKPKV 60

Qy     471 MQTCNLFDCPKWIAMEWSQCTVTCGRGLRYRVVLCINHRGEHVGGCNPQLKLHIKEECVI 530
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 MQTCNLFDCPKWIAMEWSQCTVTCGRGLRYRVVLCINHRGEHVGGCNPQLKLHIKEECVI 120

Qy     531 PIPCYKPKEKSPVEAKLPWLKQAQELEETRIATEEPTFIPEPWSACSTTCGPGVQVREVK 590
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 PIPCYKPKEKSPVEAKLPWLKQAQELEETRIATEEPTFIPEPWSACSTTCGPGVQVREVK 180

Qy     591 CRVLLTFTQTETELPEEECEGPKLPTERPCLLEACDESPASRELDIPLPEDSETTYDWEY 650
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 CRVLLTFTQTETELPEEECEGPKLPTERPCLLEACDESPASRELDIPLPEDSETTYDWEY 240

Qy     651 AGFTPCTATCLGGHQEAIAVCLHIQTQQTVNDSLCDMVHRPPAMSQACNTEPCPPRWHVG 710
           ||||||||||:|||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 AGFTPCTATCVGGHQEAIAVCLHIQTQQTVNDSLCDMVHRPPAMSQACNTEPCPPRWHVG 300

Qy     711 SWGPCSATCGVGIQTRDVYCLHPGFTPAPPEECRDEKPHALQACNQFDCPPGWHIEEWQQ 770
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 SWGPCSATCGVGIQTRDVYCLHPGFTPAPPEECRDEKPHALQACNQFDCPPGWHIEEWQQ 360

Qy     771 CSRTCGGGTQNRRVTCRQLLTDGSFLNLSDELCQGPKASSHKSCARTDCPPHLAVGDWSK 830
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 CSRTCGGGTQNRRVTCRQLLTDGSFLNLSDELCQGPKASSHKSCARTDCPPHLAVGDWSK 420

Qy     831 CSVSCGVGIQRRKQVCQRLAAKGRRIPLSEMMCRDLPGFPLVRSCQMPECSKIKSEMKTK 890
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 CSVSCGVGIQRRKQVCQRLAAKGRRIPLSEMMCRDLPGFPLVRSCQMPECSKIKSEMKTK 480

Qy     891 LGEQGPQILSVQRVYIQTREEKRINLTIGSRAYLLPNTSVIIKCPVRRFQKSLIQWEKDG 950
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 LGEQGPQILSVQRVYIQTREEKRINLTIGSRAYLLPNTSVIIKCPVRRFQKSLIQWEKDG 540

Qy     951 RCLQNSKRLGITKSGSLKIHGLAAPDIGVYRCIAGSAQETVVLKLIGTDNRLIARPALRE 1010
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 RCLQNSKRLGITKSGSLKIHGLAAPDIGVYRCIAGSAQETVVLKLIGTDNRLIARPALRE 600

Qy    1011 PMREYPGMDHSEANSLGVTWHKMRQMWNNKNDLYLDDDHISNQPFLRALLGHCSNSAGST 1070
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 PMREYPGMDHSEANSLGVTWHKMRQMWNNKNDLYLDDDHISNQPFLRALLGHCSNSAGST 660

Qy    1071 NSWELKNKQFEAAVKQGAYSMDTAQFDELIRNMSQLMETGEVSDDLASQLIYQLVAELAK 1130
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 NSWELKNKQFEAAVKQGAYSMDTAQFDELIRNMSQLMETGEVSDDLASQLIYQLVAELAK 720

Qy    1131 AQPTHMQWRGIQEETPPAAQLRGETGSVSQSSHAKNSGKLTFKPKGPVLMRQSQPPSISF 1190
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 AQPTHMQWRGIQEETPPAAQLRGETGSVSQSSHAKNSGKLTFKPKGPVLMRQSQPPSISF 780

Qy    1191 NKTINSRIGNTVYITKRTEVINILCDLITPSEATYTWTKDGTLLQPSVKIILDGTGKIQI 1250
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 NKTINSRIGNTVYITKRTEVINILCDLITPSEATYTWTKDGTLLQPSVKIILDGTGKIQI 840

Qy    1251 QNPTRKEQGIYECSVANHLGSDVESSSVLYAEAPVILSVERNITKPEHNHLSVVVGGIVE 1310
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 QNPTRKEQGIYECSVANHLGSDVESSSVLYAEAPVILSVERNITKPEHNHLSVVVGGIVE 900

Qy    1311 AALGANVTIRCPVKGVPQPNITWLKRGGSLSGNVSLLFNGSLLLQNVSLENEGTYVCIAT 1370
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 AALGANVTIRCPVKGVPQPNITWLKRGGSLSGNVSLLFNGSLLLQNVSLENEGTYVCIAT 960

Qy    1371 NALGKAVATSVLHLLERRWPESRIVFLQGHKKYILQAINTRTNSNDPTGEPPPQEPFWEP 1430
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 NALGKAVATSVLHLLERRWPESRIVFLQGHKKYILQAINTRTNSNDPTGEPPPQEPFWEP 1020

Qy    1431 GNWSHCSATCGHLGARIQRPQCVMANGQEVSEALCDHLQKPLAGFEPCNIRDCPARWFTS 1490
           ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db    1021 GNWSHCSATCGHLGARIQRPQCVMANGQEVSEALCDHLQKPLAGFEPCNIRDCPARWFTS 1080

Qy    1491 VWSQCSVSCGEGYHSRQVTCKRTKANGTVQVVSPRACAPKDRPLGRKPCEGHPCVQWEPG 1550
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 VWSQCSVSCGEGYHSRQVTCKRTKANGTVQVVSPRACAPKDRPLGRKPCEGHPCVQWEPG 1140

Qy    1551 NRCPGRCMGRAVRMQQRHTACQHNSSDSNCDDRKRPTLRRNCTSGACDVCWHTGPWKPCT 1610
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 NRCPGRCMGRAVRMQQRHTACQHNSSDSNCDDRKRPTLRRNCTSGACDVCWHTGPWKPCT 1200

Qy    1611 AACGRGFQSRRVDCIHTRSCKPVAKRHCVQKKKPISWRHCLGPSCDRDCTDTTHYCMFVK 1670
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 AACGRGFQSRRVDCIHTRSCKPVAKRHCVQKKKPISWRHCLGPSCDRDCTDTTHYCMFVK 1260

Qy    1671 HLNLCSLDRYKQRCCQSCQEG 1691
           |||||||||||||||||||||
Db    1261 HLNLCSLDRYKQRCCQSCQEG 1281


RESULT 5
AAE10608
ID   AAE10608 standard; Protein; 1281 AA.
XX
AC   AAE10608;
XX
DT   10-DEC-2001  (first entry)
XX
DE   Human novel KIAA1233-like protein, NOV-1.
XX
KW   Human; NOV-X protein; KIAA1233-like protein; STE20-like protein; tumour;
KW   trypsin inhibitor-like protein; gene therapy; haematopoietic; illness;
KW   immunological disorder; neurodegenerative disorder; Alzheimer's disease;
KW   Parkinson's disease; immunomodulatory; pharmacogenomic; haemostatic;
KW   human immunodeficiency virus; HIV; fertility disorder; neuroprotective;
KW   cytostatic; nootropic; anti-infertility; cancer; chromosome 15;
KW   NOV-1 protein.
XX
OS   Homo sapiens.
XX
PN   WO200162928-A2.
XX
PD   30-AUG-2001.
XX
PF   26-FEB-2001; 2001WO-US06151.
XX
PR   25-FEB-2000; 2000US-0184951.
PR   28-FEB-2000; 2000US-0185548.
PR   01-MAR-2000; 2000US-0185967.
PR   18-APR-2000; 2000US-0197723.
PR   27-APR-2000; 2000US-0199957.
PR   23-FEB-2001; 2001US-0789390.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Vernet CAM,  Fernandes E,  Shimkets RA,  Macdougall J,  Spaderna SK;
XX
DR   WPI; 2001-582051/65.
DR   N-PSDB; AAD17756.
XX
PT   New isolated KIAA1233-like, STE20-like, or trypsin inhibitor-like
PT   polypeptide for diagnosing and treating pathological disorders, such as
PT   Parkinson's disease and for use in pharmacogenomics
XX
PS   Claim 1; Page 7; 189pp; English.
XX
CC   The invention relates to novel human polypeptides referred as NOV-X
CC   and their corresponding nucleic acid sequences. NOV-X collectively
CC   include NOV-1, NOV-2a and NOV-2b which are novel KIAA1233-like
CC   polypeptides, NOV-3a, NOV-3b, NOV-3c and NOV-3d which are novel
CC   STE20-like polypeptides and NOV-4a, NOV-4b, NOV-4c, NOV-4d and NOV-4e
CC   which are novel trypsin inhibitor-like polypeptides. NOV-X is used to
CC   identify a potential therapeutic agent that can modulate its activity
CC   and can be used for treating a pathology related to aberrant expression
CC   or aberrant physiological interactions of NOV-X. NOV-X or its DNA is
CC   used to determine the presence or predisposition to a disease associated
CC   with altered levels of NOV-X. NOV-X, its DNA and its antibody are used
CC   to treat or prevent a pathology associated with NOV X. The pathological
CC   states that can be treated or prevented are haematopoietic, cancer,
CC   immunological, tumour, neurodegenerative (e.g. Alzheimer's and
CC   Parkinson's disease), human immunodeficiency virus (HIV) illness and
CC   fertility disorders. NOV X and its DNA are used in pharmacogenomics for
CC   predictive medicine. NOV X DNA is used in gene therapy. The present
CC   sequence is human novel KIAA1233 like protein. NOV 1. NOV 1 gene is
CC   located at chromosome 15.
XX
SQ   Sequence   1281 AA;

  Query Match             75.5%;  Score 7060;  DB 22;  Length 1281;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 1279;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        411 MPYDHFQPLPRWEHNPWIACSVSCGGGIQRRSFVCVEESMHGEILQVEEWKCMYAPKPKV 470
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MPYDHFQPLPRWEHNPWIACSVSCGGGIQRRSFVCVEESMHGEILQVEEWKCMYAPKPKV 60

Qy        471 MQTCNLFDCPKWIAMEWSQCTVTCGRGLRYRVVLCINHRGEHVGGCNPQLKLHIKEECVI 530
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 MQTCNLFDCPKWIAMEWSQCTVTCGRGLRYRVVLCINHRGEHVGGCNPQLKLHIKEECVI 120

Qy        531 PIPCYKPKEKSPVEAKLPWLKQAQELEETRIATEEPTFIPEPWSACSTTCGPGVQVREVK 590
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 PIPCYKPKEKSPVEAKLPWLKQAQELEETRIATEEPTFIPEPWSACSTTCGPGVQVREVK 180

Qy        591 CRVLLTFTQTETELPEEECEGPKLPTERPCLLEACDESPASRELDIPLPEDSETTYDWEY 650
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CRVLLTFTQTETELPEEECEGPKLPTERPCLLEACDESPASRELDIPLPEDSETTYDWEY 240

Qy        651 AGFTPCTATCLGGHQEAIAVCLHIQTQQTVNDSLCDMVHRPPAMSQACNTEPCPPRWHVG 710
              ||||||||||:|||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGFTPCTATCVGGHQEAIAVCLHIQTQQTVNDSLCDMVHRPPAMSQACNTEPCPPRWHVG 300

Qy        711 SWGPCSATCGVGIQTRDVYCLHPGETPAPPEECRDEKPHALQACNQFDCPPGWHIEEWQQ 770
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 SWGPCSATCGVGIQTRDVYCLHPGETPAPPEECRDEKPHALQACNQFDCPPGWHIEEWQQ 360

Qy        771 CSRTCGGGTQNRRVTCRQLLTDGSFLNLSDELCQGPKASSHKSCARTDCPPHLAVGDWSK 830
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CSRTCGGGTQNRRVTCRQLLTDGSFLNLSDELCQGPKASSHKSCARTDCPPHLAVGDWSK 420

Qy        831 CSVSCGVGIQRRKQVCQRLAAKGRRIPLSEMMCRDLPGEPLVRSCQMPECSKIKSEMKTK 890
              |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db        421 CSVSCGVGIQRRKQVCQRLAAKGRRIPLSEMMCRDLPGLPLVRSCQMPECSKIKSEMKTK 480

Qy        891 LGEQGPQILSVQRVYIQTREEKRINLTIGSRAYLLPNTSVIIKCPVRRFQKSLIQWEKDG 950
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 LGEQGPQILSVQRVYIQTREEKRINLTIGSRAYLLPNTSVIIKCPVRRFQKSLIQWEKDG 540

Qy        951 PCLQNSKRLGITKSGSLKIHGLAAPDIGVYRCIAGSAQETVVLKLIGTDNRLIARPALRE 1010
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 PCLQNSKRLGITKSGSLKIHGLAAPDIGVYRCIAGSAQETVVLKLIGTDNRLIARPALRE 600

Qy       1011 PMREYPGMDHSEANSLGVIWHKMRQMWNNKNDLYLDDDHISNQPFLRALLGHCSNSAGST 1070
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 PMREYPGMDHSEANSLGVIWHKMRQMWNNKNDLYLDDDHISNQPFLRALLGHCSNSAGST 660

Qy       1071 NSWELKNKQFEAAVKQGAYSMDTAQFDELIRNMSQLMETGEVSDDLASQLIYQLVAELAK 1130
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 NSWELKNKQFEAAVKQGAYSMDTAQFDELIRNMSQLMETGEVSDDLASQLIYQLVAELAK 720

Qy       1131 AQPTHMQWRGIQEETPPAAQLRGETGSVSQSSHAKNSGKLTFKPKGPVLMRQSQPPSISF 1190
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AQPTHMQWRGIQEETPPAAQLRGETGSVSQSSHAKNSGKLTFKPKGPVLMRQSQPPSISF 780

Qy       1191 NKTINGFIGNIVYTIKRTEVINLDSDLITTSEAIYIWIKNGTLLQTSVKTTLNGTGKIQI 1250
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 NKTINGFIGNIVYTIKRTEVINLDSDLITTSEAIYIWIKNGTLLQTSVKTTLNGTGKIQI 840

Qy       1251 QNPTEKEQGTYECSVANHLGSDVESSSVLYAEAPVILSVERNITKPEHNHLSVVVGGIVE 1310
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 QNPTEKEQGTYECSVANHLGSDVESSSVLYAEAPVILSVERNITKPEHNHLSVVVGGIVE 900

Qy       1311 AALGANVTIRCEVKGVPQPNITWLKRGGSLSGNVSLLFNGSLLLQNVSLENEGTYVCIAT 1370
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AALGANVTIRCEVKGVPQPNITWLKRGGSLSGNVSLLFNGSLLLQNVSLENEGTYVCIAT 960

Qy       1371 NALGKAVATSVLHLLERRWPESRIVFLQGHKKYILQATNTRTNSNDPTGEPPPQEPFWEP 1430
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 NALGKAVATSVLHLLERRWPESRIVFLQGHKKYILQATNTRTNSNDPTGEPPPQEPFWEP 1020

Qy       1431 GNWSHCSATCGHLGARIQRPQCVMANGQEVSEALCDHLQKPLAGFEPCNIRPCPARWFTS 1490
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GNWSHCSATCGHLGARIQRPQCVMANGQEVSEALCDHLQKPLAGFEPCNIRPCPARWFTS 1080

Qy       1491 VWSQCSVSCGEGYHSRQVTCKRTKANGTVQVVSPRACAPKDRPLGRKPCFGHPCVQWEPG 1550
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 VWSQCSVSCGEGYHSRQVTCKRTKANGTVQVVSPRACAPKDRPLGRKPCFGHPCVQWEPG 1140

Qy       1551 NRCPGRCMGRAVRMQQRHTACQHNSSDSNCDDRKRPTLRRNCTSGACDVCWHTGPWKPCT 1610
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 NRCPGRCMGRAVRMQQRHTACQHNSSDSNCDDRKRPTLRRNCTSGACDVCWHTGPWKPCT 1200

Qy       1611 AACGRGFQSRKVDCIHTRSCKPVAKRHCVQKKKPISWRHCLGPSCDRDCTDTTHYCMFVK 1670
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 AACGRGFQSRKVDCIHTRSCKPVAKRHCVQKKKPISWRHCLGPSCDRDCTDTTHYCMFVK 1260

Qy       1671 HLNDCSLDRYKQRCCQSCQEG 1691
              |||||||||||||||||||||
Db       1261 HLNDCSLDRYKQRCCQSCQEG 1281


RESULT 6
AAB83370
ID   AAB83370 standard; Protein; 926 AA.
XX
AC   AAB83370;
XX
DT   26-MAR-2002  (first entry)
XX
DE   NOV14 protein sequence.
XX
KW   NOV; Cytostatic; contraceptive; antiinflammatory, immunomodulatory;
KW   cardiovascular; casein kinase II phosphorylation site; contraception;
KW   serine/threonine kinase; Peutz-Jeghers syndrome; cellular proliferation;
KW   epidermal growth factor, cell development; apoptosis; cell adhesion;
KW   growth migration; cell structure; motility; cancer; immune disorder;
KW   inflammatory disorder; cellular adhesion disorder; long QT syndrome;
KW   cardiovascular disease; hypertrophic cardiomyopathy; marfan syndrome;
KW   therapy; NOV1; NOV2; NOV3; NOV4; NOV5; NOV6; NOV7; NOV8; NOV9; NOV10;
KW   NOV11, NOV12, NOV13, NOV14, NOV15; NOV16.
XX
OS   Unidentified.
XX
PN   WO200136638-A2.
XX
PD   25-MAY-2001.
XX
PF   17-NOV-2000; 2000WO-US31543.
XX
PR   19-NOV-1999;   99US-166366.
PR   29-NOV-1999;   99US-167781.
PR   08-MAR-2000; 2000US-187944.
PR   16-NOV-2000; 2000US-0715417.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Lichenstein H,  Vernet C,  Fernandes E.
XX
DR   WPI; 2001-648134/74.
DR   N-PSDB; AAF87125.
XX
PT   Novel human polypeptides and the nucleic acids that encode them useful
PT   for preventing, diagnosing and treating e.g. cancer, inflammation and
PT   immune disorders -
XX
PS   Claim 1; Page 43-45; 141pp; English.
XX
CC   This sequence is the NOV14 protein. The invention relates to
CC   the NOV1-NOV16 proteins, and their coding sequences. The proteins have
CC   cytostatic; contraceptive; antiinflammatory; immunomodulatory; and
CC   cardiovascular activities. The sequences may be used in the prevention,
CC   diagnosis and treatment of diseases associated with inappropriate NOVX
CC   expression. They may be used to treat disorders associated with decreased
CC   expression by rectifying mutations or deletions in a patient's genome
CC   that affect the activity of protein by expressing inactive proteins or to
CC   supplement the patients own production of protein. They are used to
CC   produce NOVX proteins, by inserting the nucleic acid into a cell and
CC   culturing it to express the protein. The DNA may be used as DNA probes in
CC   assays to detect and quantitate the presence of similar DNAs in samples,
CC   and which patients may need restorative therapy. The NOVX protein may
CC   also be used as antigens in the production of antibodies (Abs) against
CC   NOVX and in assays to identify modulators of NOVX expression and
CC   activity. The anti-NOVX Abs and antagonist are used to down regulate
CC   expression and activity. The anti-NOVX Abs are used for detecting the
CC   presence of NOVX in samples. Disorders that may be prevented, diagnosed
CC   and/or treated vary depending on the NOVX protein. NOV1, NOV3, NOV5,
CC   NOV7, NOV9-11 and NOV13-16 have casein kinase II phosphorylation sites
CC   characteristic of serine/threonine kinases, and are used to treat
CC   kinase-related disorders (e.g. Peutz-Jeghers syndrome, cellular
CC   proliferation and contraception). NOV2-3, NOV6 and NOV8 are homologous to
CC   the epidermal growth factor (EGF)-like super family and are involved in,
CC   e.g. regulation of cell development, apoptosis, cell adhesion, growth
CC   migration, cell structure and motility and protein management, and are
CC   used to treat cancers, inflammatory disorders, immune disorders and
CC   cellular adhesion disorders. NOV6-10 are homologous to EGF-like
CC   fibrillin proteins and are used to treat cardiovascular disease e.g.
CC   hypertrophic cardiomyopathy, long-QT syndrome and marfan syndrome.
XX
SQ   Sequence   926 AA;

  Query Match             54.4%;  Score 5086;  DB 22;  Length 926;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches  925;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        314 MADFIFKIRYTAAKDSVVQFFFYQPISHQWRQTDFPPCTVTCGGGYQLNSAECVDIRLKR 373
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MADFIFKIRYTAAKDSVVQFFFYQPISHQWRQTDFPPCTVTCGGGYQLNSAECVDIRLKR 60

Qy        374 VVPDHYCHYYPENVKPKPKLKECSMDPCPSSDGFKEIMPYDHPQPLPEWEHNPWTACSVS 433
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VVPDHYCHYYPENVKPKPKLKECSMDPCPSSDGFKEIMPYDHPQPLPEWEHNPWTACSVS 120

Qy        434 CGGGIQRRSFVCVEESMHGEILQVEEWKCMYAPKPKVMQTCNLFDCPKWIAMEWSQCTVT 493
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CGGGIQRRSFVCVEESMHGEILQVEEWKCMYAPKPKVMQTCNLFDCPKWIAMEWSQCTVT 180

Qy        494 CGRGLRYRVVLCINHRGEHVGGCNPQLKLHIKEECVIPIPCYKPKEKSPVEAKLPWLKQA 553
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CGRGLRYRVVLCINHRGEHVGGCNPQLKLHIKEECVIPIPCYKPKEKSPVEAKLPWLKQA 240

Qy        554 QELEETRIATEEPTFIPEPWSACSTTCGPGVQVREVKCRVLLTPTQTETELPEEECEGPK 613
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 QELEETRIATEEPTFIPEPWSACSTTCGPGVQVREVKCRVLLTPTQTETELPEEECEGPK 300

Qy        614 LPTERPCLLEACDESPASKELDIPLPEDSETTYDWEYAGFTPCTATCLGGHQEAIAVCLH 673
              |||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||
Db        301 LPTERPCLLEACDESPASKELDIPLPEDSETTYDWEYAGFTPCTATCVGGHQEAIAVCLH 360

Qy        674 IQTQQTVNDSLCDMVHRPPAMSQACNTEPCPPRWHVGSWGPCSATCGVGIQTRDVYCLHP 733
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 IQTQQTVNDSLCDMVHRPPAMSQACNTEPCPPRWHVGSWGPCSATCGVGIQTRDVYCLHP 420

Qy        734 GETPAPPEECRDEKPHALQACNQFDCPPGWHIEEWQQCSRTCGGGTQNRRVTCRQLLTDG 793
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GETPAPPEECRDEKPHALQACNQFDCPPGWHIEEWQQCSRTCGGGTQNRRVTCRQLLTDG 480

Qy        794 SFLNLSDELCQGPKASSHKSCARTDCPPHLAVGDWSKCSVSCGVGIQRRKQVCQRLAAKG 853
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 SFLNLSDELCQGPKASSHKSCARTDCPPHLAVGDWSKCSVSCGVGIQRRKQVCQRLAAKG 540

Qy        854 RRIFLSEMMCFDLPGFHIVESCQMFECRFRSEMFINIEGQGFQIISVQFVYIQIPLEKR 913
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 RRIFLSEMMCFDLPGFHIVESCQMFECRFRSEMFINIEGQGFQIISVQFVYIQIPLEKR 600

Qy        914 INLTIGSRAYLLPNTSVIIKCPVRRFQRSLIQWFFDGRCLQNSFRLCIIKSGISLKIHCLA 973
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 INLTIGSRAYLLPNTSVIIKCPVRRFQRSLIQWFFDGRCLQNSFRLCIIKSGISLKIHCLA 660

Qy        974 APDIGVYRCIAGSAQETVVLKLIGTDNRLIARPALREPMREYPGMDHSEANSLCVIWHKM 1033
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 APDIGVYRCIAGSAQETVVLKLIGTDNRLIARPALREPMREYPGMDHSEANSLCVIWHKM 720

Qy       1034 KQMWNNKNLLYLLGDHISNQFFLKALLGHCSNSAGSINSWLLKNKQFLAAVKQGAYSMDI 1093
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 KQMWNNKNLLYLLGDHISNQFFLKALLGHCSNSAGSINSWLLKNKQFLAAVKQGAYSMDI 780

Qy       1094 AQFDELIRNMSQLMETGEVSDDLASQLIYQLVAELAKAQPTHMQWRGIQEETPPAAQLRG 1153
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 AQFDELIRNMSQLMETGEVSDDLASQLIYQLVAELAKAQPTHMQWRGIQEETPPAAQLRG 840

Qy       1154 ETGSVSQSSHAKNSCKLTEKPFKEVLMSQSQFEPISENKLLNSRLQNIVFLLKKILVINI 1213
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ETGSVSQSSHAKNSCKLTEKPFKEVLMSQSQFEPISENKLLNSRLQNIVFLLKKILVINI 900

Qy       1214 LCDLITPSEATYTWTKDGTLLQPSVK 1239
              ||||||||||||||||||||||||||
Db        901 LCDLITPSEATYTWTKDGTLLQPSVK 926


RESULT 7
AAB40500
ID   AAB40500 standard; Protein; 886 AA.
XX
AC   AAB40500;
XX
DT   08-FEB-2001  (first entry)
XX
DE   Human ORFX ORF264 polypeptide sequence SEQ ID NO:528.
XX
KW   Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
KW   vulnerary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
KW   anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KW   immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW   hypotensive; dermatological; immunosuppressive; antiinflammatory;
KW   antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
KW   antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
KW   neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW   cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW   cholesterol ester storage; systemic lupus erythematosus; infection;
KW   severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW   allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW   bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW   thrombosis; contraceptive.
XX
OS   Homo sapiens.
XX
PN   WO200058473-A2.
XX
PD   05-OCT-2000.
XX
PF   31-MAR-2000; 2000WO-US08621.
XX
PR   31-MAR-1999;   99US-0127607.
PR   02-APR-1999;   99US-0127636.
PR   05-APR-1999;   99US-0127728.
PR   30-MAR-2000; 2000US-0540763.
XX
PA   (CURA-) CURAGEN CORP.
XX
```

```
PI   Shimkets RA, Leach M.
XX
DR   WPI; 2000-682362/67.
DR   N-PSDB; AAC74709.
XX
PT   Novel nucleic acids and peptides derived from open reading frame X,
PT   useful for treating e.g. cancers, proliferative disorders,
PT   neurodegenerative disorders and cardiovascular disease -
XX
PS   Claim 11; Page 693-695; 5507pp; English.
XX
CC   AAC74446 to AAC77636 encode the proteins given in AAB40237 to AAB43397,
CC   which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC   sequences have activities such as: cytostatic; hepatotropic; vulnerary;
CC   antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
CC   osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
CC   immunostimulant, cardiant, thrombolytic; coagulant; vasotropic;
CC   antidiabetic; hypotensive; dermatological; immunosuppressive;
CC   antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
CC   antithyroid; and antianaemic. The sequences can be used for determining
CC   the presence of or predisposition to, or preventing or treating
CC   pathological conditions associated with an ORFX-associated disorder. The
CC   nucleic acids can be used to express ORFX proteins in gene therapy
CC   vectors. The proteins and nucleic acids may be used to treat cancers,
CC   proliferative disorders, neurodegenerative disorders, osteoarthritis,
CC   graft vs host disease, cardiovascular disease, diabetes mellitus,
CC   hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
CC   erythematosus, severe combined immunodeficiency (SCID), AIDS, viral
CC   bacterial or fungal infection, malaria, autoimmune disorders, asthma,
CC   allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
CC   nocturnal haemoglobinuria, antiinflammatory disease; to enhance
CC   coagulation; to inhibit thrombosis; and as a contraceptive.
XX
SQ   Sequence   886 AA;

  Query Match             50.2%;  Score 4693;  DB 21;  Length 886;
  Best Local Similarity   99.2%;  Pred. No. 1.1e-299;
  Matches  876;  Conservative    0;  Mismatches    5;  Indels    2;  Gaps    1;

Qy        591 CRVLLIFTQIEIELPEEEVEGPKLPIERPCLLEACDESPASRELDIPLPEDSEITYDWEY 650
              |||||||||||||||||| ||||||||||| ||||||||||||||||||||||| |||||
Db          2 CRVLLIFTQIEIELPEEECEGPKLPIERPCFLEACDESPASRELDIPLPEDSFIAYDWEY 61

Qy        651 AGFIPCIAICDGHQEAIAVCLHIQTQQTVNDSLCDMVHRPPAMSQACNTEPCPPRWHVG 710
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 AGFIPCIAICDGHQEAIAVCLHIQTQQTVNDSLCDMVHRPPAMSQACNTEPCPPRWHVG 121

Qy        711 SWGPCSATCGVGIQTRDVYCLHPGETPAPPEECRDEKPHALQACNQFDCPPGWHIEEWQQ 770
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 SWGPCSATCGVGIQTRDVYCLHPGETPAPPEECRDEKPHALQACNQFDCPPGWHIEEWQQ 181

Qy        771 CSRTCGGGTQNRRVTCRQLLTDGSFLNLSDELCQGPKASSHKSCARTDCPPHLAVGDWSK 830
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 CSRTCGGGTQNRRVTCRQLLTDGSFLNLSDELCQGPKASSHKSCARTDCPPHLAVGDWSK 241

Qy        831 CSVSCGVGIQRRKQVCQRLAAKGRRIPLSEMMCRDLPGLPLVRSCQMPECSKIKSEMKTK 890
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 CSVSCGVGIQRRKQVCQRLAAKGRRIPLSEMMCRDLPGLPLVRSCQMPECSKIKSEMKTK 301

Qy        891 LGEQGPQILSVQRVYIQTREEKRINLTIGSRAYLLPNTSVIIKCPVRRFQKSLIQWEKDG 950
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 LGEQGPQILSVQRVYIQTREEKRINLTIGSRAYLLPNTSVIIKCPVRRFQKSLIQWEKDG 361

Qy        951 RCLQNSKRLGITKSGSLKIHGLAAPDIGVYRCIAGSAQETVVLKLIGTDNRLIARPALRE 1010
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 RCLQNSKRLGITKSGSLKIHGLAAPDIGVYRCIAGSAQETVVLKLIGTDNRLIARPALRE 421

Qy       1011 PMREYPGMDHSEANSLGVTWHKMRQMWNNKNDLYLDDDHISNQPFLRALLGHCSNSAGST 1070
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 PMREYPGMDHSEANSLGVTWHKMRQMWNNKNDLYLDDDHISNQPFLRALLGHCSNSAGST 481

Qy       1071 NSWELKNKQFEAAVKQGAYSMDTAQFDELIRNMSQLMETGEVSDDLASQLIYQLVAELAK 1130
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 NSWELKNKQFEAAVKQGAYSMDTAQFDELIRNMSQLMETGEVSDDLASQLIYQLVAELAK 541

Qy       1131 AQPTHMQWRGIQEETPPAAQLRGETGSVSQSSHAKNSGKLTFKPKGPVLMRQSQPPSISF 1190
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        542 AQPTHMQWRGIQEETPPAAQLRGETGSVSQSSHAKNSGKLTFKPKGPVLMRQSQPPSISF 601

Qy       1191 NKTINSRIGNTVYITKRTEVINILCDLITPSEATYTWTKDGTLLQPSVKIIIDGTGKIQI 1250
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        602 NKTINSRIGNTVYITKRTEVINILCDLITPSEATYTWTKDGTLLQPSVKIIIDGTGKIQI 661

Qy       1251 QNPTRKEQGIYECSVANHLGSDVESSSVLYAEAPVILSVERN-IKPEHNHLSVVVGGIVE 1310
              |||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db        662 QNPTRKEQGIYECSVANHLGSDVESSSVLYAEAPVILSVERNITKPEHNHLSVVVGGIVE 721

Qy       1311 AALGANVTIRCPVKGVPQPNITWLKRGGSLSGNVSLLFNGSLLLQNVSLENEGTYVCIAT 1370
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        722 AALGANVTIRCPVKGVPQPNITWLKRGGSLSGNVSLLFNGSLLLQNVSLENEGTYVCIAT 781

Qy       1371 NALGKAVATSVLHLLERRWPESRIVFLQGHKKYILQATNTRTNSNDPTGEPPPQEPFWEP 1430
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        782 NALGKAVATSVLHLLERRWPESRIVFLQGHKKYILQATNTRTNSNDPTGEPPPQEPFWEP 841

Qy       1431 GNWSHCSATCGHLGARIQRPQCVMANGQEVSEALCDHLQKPLA 1473
              ||||||||||||||||||||||||||||||||||||   |  |
Db        842 GNWSHCSATCGHLGARIQRPQCVMANGQEVSEALCD--QPPEA 882


RESULT 8
AAE07868
ID   AAE07868 standard; Protein; 845 AA.
XX
AC   AAE07868;
XX
DT   01 NOV 2001  (first entry)
XX
DE   Novel human protein (NHP) #6.
XX
KW   Novel human protein; NHP; thrombospondin; gene therapy; cancer;
KW   cytostatic; antisense therapy; angiogenesis; biological disorder.
XX
OS   Homo sapiens.
XX
PN   WO200161011-A2.
XX
PD   23-AUG-2001.
XX
PF   15-FEB-2001; 2001WO-US05240.
XX
PR   17-FEB-2000; 2000US-0183282.
XX
PA   (LEXI-) LEXICON GENETICS INC.
XX
PI   Donoho G,  Scoville J,  Turner CA,  Friedrich G,  Zambrowicz B;
PI   Sands AT;
XX
DR   WPI; 2001-511759/56.
DR   N-PSDB; AAD14369.
XX
PT   An isolated nucleic acid encoding a thrombospondin useful as a
PT   hybridization probe and gene therapy treatments of cancer
XX
PS   Disclosure; Page 43-45; 56pp; English.
XX
CC   The invention relates to novel human polynucleotides encoding proteins
CC   that share sequence similarity with animal proteins having thrombospondin
CC   repeats. NHPs are expressed in human cell lines, pituitary, lymph nodes,
CC   prostrate, testis, adrenal gland, uterus, foetal kidney, foetal lung and
CC   gene trapped human cells. NHPs are thrombospondins useful for treating
CC   biological disorders involving angiogenesis, cancer and development and
CC   also in pharmacogenomic applications. NHPs are useful as a hybridisation
CC   probe for screening libraries, assessing gene expression patterns and
CC   also in gene therapy. Proteins having thrombospondin repeats act as
CC   receptors, secreted extracellular matrix proteins and proteases. The
CC   present sequence is novel human protein (NHP).
XX
SQ   Sequence   845 AA;

  Query Match              50.2%;  Score 4689;  DB 22;  Length 845;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-299;
  Matches  830;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MASWISPWWVLIGMVFMHSPLPQTTAEKSPGAYFLPEFALSPQGSFLEDTTGEQFLTYRY 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MASWISPWWVLIGMVFMHSPLPQTTAEKSPGAYFLPEFALSPQGSFLEDTTGEQFLTYRY 60

Qy         61 DDQTSRNTRSDEDKDGNWDAWGDWSDCSRTCGGGASYSLRRCLTCRNCEGQNIRYKTCSN 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DDQTSRNTRSDEDKDGNWDAWGDWSDCSRTCGGGASYSLRRCLTCRNCEGQNIRYKTCSN 120

Qy        121 HDCPPDAEDFRAQQCSAYNDVQYQGHYYEWLPRYNDPAAPCALKCHAQGQNLVVELAPKV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 HDCPPDAEDFRAQQCSAYNDVQYQGHYYEWLPRYNDPAAPCALKCHAQGQNLVVELAPKV 180

Qy        181 LDGTRCNTDSLDMCISGICQAVGCDPQLGSNAKEDNCGVCAGDGSTCRLVRGQSKSHVSP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 LDGTRCNTDSLDMCISGICQAVGCDPQLGSNAKEDNCGVCAGDGSTCRLVRGQSKSHVSP 240

Qy        241 EKREENVIAVPLGSRSVRITVKGPAHLFIESKTLQGSKGEHSFNSPGVFVVENTTVEFQR 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 EKREENVIAVPLGSRSVRITVKGPAHLFIESKTLQGSKGEHSFNSPGVFVVENTTVEFQR 300

Qy        301 GSERQTFKIPGPLMADFIFKTRYTAAKDSVVQFFFYQPISHQWRQTDFFPCTVTCGGGYQ 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GSERQTFKIPGPLMADFIFKTRYTAAKDSVVQFFFYQPISHQWRQTDFFPCTVTCGGGYQ 360

Qy        361 LNSAECVDIRLKRVVPDHYCHYYPENVKPKPKLKECSMDPCPSSDGFKEIMPYDHFQPLP 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LNSAECVDIRLKRVVPDHYCHYYPENVKPKPKLKECSMDPCPSSDGFKEIMPYDHFQPLP 420

Qy        421 RWEHNPWTACSVSCGGGIQRRSFVCVEESMHGEILQVEEWKCMYAPKPKVMQTCNLFDCP 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 RWEHNPWTACSVSCGGGIQRRSFVCVEESMHGEILQVEEWKCMYAPKPKVMQTCNLFDCP 480

Qy        481 KWIAMEWSQCTVTCGRGLRYRVVLCINHRGEHVGGCNPQLKLHIKEECVIPIPCYKPKEK 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 KWIAMEWSQCTVTCGRGLRYRVVLCINHRGEHVGGCNPQLKLHIKEECVIPIPCYKPKEK 540

Qy        541 SPVEAKLPWLKQAQELEETRIATEEPTFIPEPWSACSTTCGPGVQVREVKCRVLLTFTQT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 SPVEAKLPWLKQAQELEETRIATEEPTFIPEPWSACSTTCGPGVQVREVKCRVLLTFTQT 600

Qy        601 ETELPEEECEGPKLPTERPCLLEACDESPASRELDIPLPEDSETTYDWEYAGFTPCTATC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ETELPEEECEGPKLPTERPCLLEACDESPASRELDIPLPEDSETTYDWEYAGFTPCTATC 660

Qy        661 LGGHQEAIAVCLHIQTQQTVNDSLCDMVHRPPAMSQACNTEPCPPRWHVGSWGPCSATCG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 LGGHQEAIAVCLHIQTQQTVNDSLCDMVHRPPAMSQACNTEPCPPRWHVGSWGPCSATCG 720

Qy        721 VGIQTRDVYCLHPGETPAPPEECRDEKPHALQACNQFDCPPGWHIEEWQQCSRTCGGGTQ 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 VGIQTRDVYCLHPGETPAPPEECRDEKPHALQACNQFDCPPGWHIEEWQQCSRTCGGGTQ 780

Qy        781 NRRVTCRQLLTDGSFLNLSDELCQGPKASSHKSCARTDCPPHLAVGDWSF 830
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 NRRVTCRQLLTDGSFLNLSDELCQGPKASSHKSCARTDCPPHLAVGDWSF 830


RESULT 9
AAE07869
ID   AAE07869 standard; Protein; 771 AA.
XX
AC   AAE07869;
XX
DT   01-NOV-2001  (first entry)
XX
DE   Novel human protein (NHP) #7.
XX
KW   Novel human protein; NHP; thrombospondin; gene therapy; cancer;
KW   cytostatic, antisense therapy, angiogenesis, biological disorder.
XX
OS   Homo sapiens.
XX
PN   WO200161011-A2.
XX
PD   23-AUG-2001.
XX
PF   15-FEB-2001; 2001WO-US05290.
XX
PR   17 FEB 2000; 2000US 0183282.
XX
PA   (LEXI-) LEXICON GENETICS INC.
XX
PI   Donoho G,  Scoville J,  Turner CA,  Friedrich G,  Zambrowicz B;
PI   Sands AT;
XX
DR   WPI; 2001-514776/56.
DR   N-PSDB; AAD14370.
XX
PT   An isolated nucleic acid encoding a thrombospondin useful as a
PT   hybridization probe and gene therapy treatments of cancer -
XX
PS   Disclosure; Page 46-48; 56pp; English.
XX
CC   The invention relates to novel human polynucleotides encoding proteins
CC   that share sequence similarity with animal proteins having thrombospondin
CC   repeats. NHPs are expressed in human cell lines, pituitary, lymph nodes,
CC   prostrate, testis, adrenal gland, uterus, foetal kidney, foetal lung and
CC   gene trapped human cells. NHPs are thrombospondins useful for treating
CC   biological disorders involving angiogenesis, cancer and development and
CC   also in pharmacogenomic applications. NHPs are useful as a hybridisation
CC   probe for screening libraries, assessing gene expression patterns and
CC   also in gene therapy. Proteins having thrombospondin repeats act as
CC   receptors, secreted extracellular matrix proteins and proteases. The
CC   present sequence is novel human protein (NHP).
XX
SQ   Sequence   771 AA;

  Query Match              45.2%;  Score 4225;  DB 22;  Length 771;
  Best Local Similarity    91.1%;  Pred. No. 5.4e-269;
  Matches  756;  Conservative    0;  Mismatches    0;  Indels   74;  Gaps    1;

Qy          1 MASWTSPWWVLIGMVFMHSPLPQTTAEKSPGAYFLPEFALSPQGSFLEDTTGEQFLTYRY 60
              ||||||||||||||||||||||||||||||
Db          1 MASWTSPWWVLIGMVFMHSPLPQTTAEKSP------------------------------ 30

Qy         61 DDQTSRNTRSDEDKDGNWDAWGDWSDCSRTCGGGASYSLRRCLTCRNCEGQNIRYKTCSN 120
                                                          ||||||||||||||||
Db         31 --------------------------------------------CRNCEGQNIRYKTCSN 46

Qy        121 HDCPPDAEDFRAQQCSAYNDVQYQGHYYEWLPRYNDPAAPCALKCHAQGQNLVVELAPKV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         47 HDCPPDAEDFRAQQCSAYNDVQYQGHYYEWLPRYNDPAAPCALKCHAQGQNLVVELAPKV 106

Qy        181 LDGTRCNTDSLDMCISGICQAVGCDPQLGSNAKEDNCGVCAGDGSTCRLVRGQSKSHVSP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 LDGTRCNTDSLDMCISGICQAVGCDPQLGSNAKEDNCGVCAGDGSTCRLVRGQSKSHVSP 166

Qy        241 EKREENVIAVPLGSRSVRITVKGPAHLFIESKTLQGSKGEHSFNSPGVFVVENTTVEFQR 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 EKREENVIAVPLGSRSVRITVKGPAHLFIESKTLQGSKGEHSFNSPGVFVVENTTVEFQR 226

Qy        301 GSERQTFKIPGPLMADFIFKTRYTAAKDSVVQFFFYQPISHQWRQTDFFPCTVTCGGGYQ 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 GSERQTFKIPGPLMADFIFKTRYTAAKDSVVQFFFYQPISHQWRQTDFFPCTVTCGGGYQ 286

Qy        361 LNSAECVDIRLKRVVPDHYCHYYPENVKPKPKLKECSMDPCPSSDGFKEIMPYDHFQPLP 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 LNSAECVDIRLKRVVPDHYCHYYPENVKPKPKLKECSMDPCPSSDGFKEIMPYDHFQPLP 346
```

```
Qy        421 RWEHNPWTACSVSCGGGIQRRSFVCVEESMHGEILQVEEWKCMYAPKPKVMQTCNLFDCP 480
              ||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||||
Db        347 RWEHNPWTACSVSCGGGIQRRSFVCVEESMHGEILQVEEWKCMYAPKPKVMQTCNLFDCP 406

Qy        481 KWIAMEWSQCTVTCGRGLRYRVVLCINHRGEHVGGCNPQLKLHIKEECVIPIPCYKPKEK 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        407 KWIAMEWSQCTVTCGRGLRYRVVLCINHRGEHVGGCNPQLKLHIKEECVIPIPCYKPKEK 466

Qy        541 SPVEAKLPWLKQAQELEETRIATEEPTFIPEPWSACSTTCGPGVQVREVKCRVLLTFTQT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        467 SPVEAKLPWLKQAQELEETRIATEEPTFIPEPWSACSTTCGPGVQVREVKCRVLLTFTQT 526

Qy        601 ETELPEEECEGPKLPTERPCLLEACDESPASRELDIPLPEDSETTYDWEYAGFTPCTATC 660
              ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        527 ETELPEEECEGPKLPTERPCLLEACDESPASRELDIPLPEDSETTYDWEYAGFTPCTATC 586

Qy        661 LGGHQEATAVCLHIQTQQTVNDSLCDMVHRPPAMSQACNTEPCPPRWHVGSWGPCSATCG 720
              |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||| ||
Db        587 LGGHQEATAVCLHIQTQQTVNDSLCDMVHRPPAMSQACNTEPCPPRWHVGSWGPCSATCG 646

Qy        721 VGTQTRDVYCLHPSGTTAPPEDCKDEKPHALQACNQFDCPPGWHIEEWQQCSRTCGGGTQ 780
              |||||||||||| |||||| ||||||||||||||||||||||||||||||||||||||||
Db        647 VGTQTRSVYCLHPSGTTAPPEDCKDEKPHALQACNQFDCPPGWHIEEWQQCSRTCGGGTQ 706

Qy        781 NRRVTCRQLLTDGSFLNLSDELCQGPKASSHKSCARTDCPPHLAVGDWSK 830
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        707 NRRVTCRQLLTDGSFLNLSDELCQGPKASSHKSCARTDCPPHLAVGDWSK 756


RESULT 10
AAE07866
ID   AAE07866 standard; Protein; 724 AA.
XX
AC   AAE07866;
XX
DT   01-NOV-2001  (first entry)
XX
DE   Novel human protein (NHP) #4.
XX
KW   Novel human protein; NHP; thrombospondin; gene therapy; cancer;
KW   cytostatic; antisense therapy; angiogenesis; biological disorder.
XX
OS   Homo sapiens.
XX
PN   WO200164011-A2.
XX
PD   23-AUG-2001.
XX
PF   15-FEB-2001; 2001WO-US05200.
XX
PR   17-FEB-2000; 2000US-0183286.
XX
PA   (LEXI-) LEXICON GENETICS INC.
XX
PI   Donoho G,  Scoville J,  Turner CA,  Friedrich G,  Zambrowicz B;
PI   Sands AT;
XX
DR   WPI; 2001-514776/56.
DR   N-PSDB; AAD14367.
XX
PT   An isolated nucleic acid encoding a thrombospondin useful as a
PT   hybridization probe and gene therapy treatments of cancer
XX
PS   Disclosure; Page 38-40; 56pp; English.
XX
CC   The invention relates to novel human polynucleotides encoding proteins
CC   that share sequence similarity with animal proteins having thrombospondin
CC   repeats. NHPs are expressed in human cell lines, pituitary, lymph nodes,
CC   prostrate, testis, adrenal gland, uterus, foetal kidney, foetal lung and
CC   gene trapped human cells. NHPs are thrombospondins useful for treating
CC   biological disorders involving angiogenesis, cancer and development and
CC   also in pharmacogenomic applications. NHPs are useful as a hybridisation
CC   probe for screening libraries, assessing gene expression patterns and
CC   also in gene therapy. Proteins having thrombospondin repeats act as
CC   receptors, secreted extracellular matrix proteins and proteases. The
CC   present sequence is novel human protein (NHP).
XX
SQ   Sequence   724 AA;

  Query Match             42.9%;  Score 3956;  DB 22;  Length 724;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-251;
  Matches  706;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MASWISPWWVLIGMVFMHSPLPQTTAEKSPGAYFLPEFALSPQGSFLEDTTGEQFLTYRY 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MASWISPWWVLIGMVFMHSPLPQTTAEKSPGAYFLPEFALSPQGSFLEDTTGEQFLTYRY 60

Qy         61 HDQTSRNTRSDLDKDGNWDAWGDWSDCSRTCGGGASYSLRRCLTGRNCEGQNIRYKTCSN 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 HDQTSRNTRSDLDKDGNWDAWGDWSDCSRTCGGGASYSLRRCLTGRNCEGQNIRYKTCSN 120

Qy        121 HDCPPDAEDFRAQQCSAYNDVQYQGHYYEWLPRYNDPAAPCALKCHAQGQNLVVELAPKV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 HDCPPDAEDFRAQQCSAYNDVQYQGHYYEWLPRYNDPAAPCALKCHAQGQNLVVELAPKV 180

Qy        181 LDGTRCNTDSLDMCISGICQAVGCDRQLGSNAKEDNCGVCAGDGSTCRLVRGQSKSHVSP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 LDGTRCNTDSLDMCISGICQAVGCDRQLGSNAKEDNCGVCAGDGSTCRLVRGQSKSHVSP 240

Qy        241 EKREENVIAVPLGSRSVRITVKGPAHLFIESKTLQGSKGEHSFNSPGVFVVENTTVEFQR 300
              ||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db        241 EKREENVIAVPLGSRSVRITVKGPAHLFIESKTLQGSKGEHSFNSPGVFVVENTTVEFQR 300

Qy        301 GSERQTFKIPGPLMADFIFKTRYTAAKDSVVQFFFYQPISHQWRQTDFFPCTVTCGGGYQ 360
              |||||||||||||||||||| |||||||||||||||||||||||||||||| ||||||||
Db        301 GSERQTFKIPGPLMADFIFKTRYTAAKDSVVQFFFYQPISHQWRQTDFFPCTVTCGGGYQ 360

Qy        361 LNSAECVDIRLKRVVPDHYCHYYPENVKPKPKLKECSMDPCPSSDGFKEIMPYDHFQPLP 420
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        361 LNSAECVDIRLKRVVPDHYCHYYPENVKPKPKLKECSMDPCPSSDGFKEIMPYDHFQPLP 420

Qy        421 RWEHNPWTACSVSCGGGIQRRSFVCVEESMHGEILQVEEWKCMYAPKPKVMQTCNLFDCP 480
              |||||||||||||||||||||||||||| ||||||||||||||||||||||||| |||||
Db        421 RWEHNPWTACSVSCGGGIQRRSFVCVEESMHGEILQVEEWKCMYAPKPKVMQTCNLFDCP 480

Qy        481 KWIAMEWSQCTVTCGRGLRYRVVLCINHRGEHVGGCNPQLKLHIKEECVIPIPCYKPKEK 540
              |||||||||||||||||| ||||||||||||||||||||||||||||||||| ||||||||
Db        481 KWIAMEWSQCTVTCGRGLRYRVVLCINHRGEHVGGCNPQLKLHIKEECVIPIPCYKPKEK 540

Qy        541 SPVEAKLPWLKQAQELEETRIATEEPTFIPEPWSACSTTCGPGVQVREVKCRVLLTFTQT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 SPVEAKLPWLKQAQELEETRIATEEPTFIPEPWSACSTTCGPGVQVREVKCRVLLTFTQT 600

Qy        601 ETELPEEECEGPKLPTERPCLLEACDESPASRELDIPLPEDSETTYDWEYAGFTPCTATC 660
              |||||||||||||||||||| |||||||||||||| |  |||||||||||||||||||||
Db        601 ETELPEEECEGPKLPTERPCLLEACDESPASRELDIPLPEDSETTYDWEYAGFTPCTATC 660

Qy        661 LGGHQEATAVCLHIQTQQTVNDSLCDMVHRPPAMSQACNTEPCPPR 706
              |||||||||| |||||||||||||||||||||||| || | ||||||
Db        661 LGGHQEATAVCLHIQTQQTVNDSLCDMVHRPPAMSQACNTEPCPPR 706


RESULT 11
AAB90551
ID   AAB90551 standard; Protein; 1749 AA.
XX
AC   AAB90551;
XX
DT   01-JUN-2001  (first entry)
XX
DE   Human secreted protein, SEQ ID NO: 89.
XX
KW   Human; secreted protein; immunomodulatory; antisclerotic;
KW   dermatological; antiinflammatory; anti-HIV; cytostatic; cardiant;
KW   vascular; anti-angiogenic; ophthalmological; neuroprotectant;
KW   nootropic; anticonvulsant; antialzheimers; antiparkinsonian;
KW   antimicrobial; vulnerary; vaccine; gene therapy; cancer;
KW   protein coordinate data; infection.
XX
OS   Homo sapiens.
XX
PN   WO200121658 A1.
XX
PD   29-MAR-2001.
XX
PF   22-SEP-2000; 2000WO-US26013.
XX
PR   24-SEP-1999;   99US-0155709.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Ni J,  Baker KP,  Birse CE,  Ebner R,  Fiscella M,  Komatsoulis GA,
PI   Lafleur DW,  Moore PA,  Olsen HS,  Rosen CA,  Ruben SA,  Soppet DR;
PI   Young PE,  Wei P,  Florence KA;
XX
DR   WPI; 2001-235311/24
DR   N-PSDB; AAF97891.
XX
PT   Nucleic acids encoding 32 human secreted polypeptides, useful for
PT   preventing, diagnosing and/or treating e.g. cancers, Parkinson's
PT   disease and diabetic retinopathy -
XX
PS   Claim 11; Fig 4; 890pp; English.
XX
CC   The present sequence is one of 32 novel human secreted polypeptides. The
CC   nucleic acid molecules and polypeptides may be used in the prevention
CC   diagnosis and treatment of diseases such as immune disorders (e.g.
CC   multiple sclerosis, systemic lupus erythematosus and human
CC   immuno-deficiency virus (HIV) infections), hyperproliferative disorders
CC   (e.g. cancers and Gaucher's disease), cardiovascular diseases
CC   (e.g. Scimitar syndrome, Chaga's cardiomyopathy and coronary
CC   arteriosclerosis), angiogenic disorders (e.g. corneal graft
CC   neovascularisation and diabetic retinopathy), neurological disorders
CC   (e.g. Huntington's chorea, Alzheimer's disease and Parkinson's disease),
CC   infectious diseases and/or for promoting wound healing, regeneration
CC   and/or chemotaxis. The nucleic acid molecules may be used to produce the
CC   secreted polypeptides. They may also be used as DNA probes in diagnostic
CC   assays to detect and quantitate the presence of similar nucleic acid
CC   sequences in samples. The polypeptides may be used as antigens in the
CC   production of antibodies and in assays to identify modulators of
CC   their expression and activity.
XX
SQ   Sequence   1745 AA;

  Query Match             41.3%;  Score 3861;  DB 22;  Length 1745;
  Best Local Similarity   42.6%;  Pred. No. 1.6e-244;
  Matches  756;  Conservative  263;  Mismatches  552;  Indels  204;  Gaps   32;

Qy         64 TSRNTRSDEDKDGNWDAWGDWSDCSRTCGGGASYSLRRCLTGRNCEGQNIRYKTCSNHDC 123
              :||   ||:||:|| :||||  ||:|||||||||||||||||||: ::|||:||||:|||| ||
Db         22 SSKTARSEEDRDGLWDAWGPWSECSRTCGGGASYSLRRCLSSKSCEGRNIRYRTCSNVDC 81

Qy        124 PPDAEDFRAQQCSAYNDVQYQGHYYEWLPRYNDPAAPCALKCHAQGQNLVVELAPKVLDG 183
              ||:| ||||||||||:||:: | :|||||   |||   ||:||| |:|   ||||||||||||
Db         82 PPEAGDFRAQQCSAHNDVKHHGQFYEWLPVSNDPDNPCSLKCQAKGTTLVVELAPKVLDG 141

Qy        184 TRCNTDSLDMCISGICQAVGCDRQLGSNAKEDNCGVCAGDGSTCRLVRGQSKSHVSPEKR 243
              |||  |:||||||||||:||  |:|||  ||||    ||||||||  |||||||||||| ||  :|   |
Db        142 TRCYTESLDMCISGLCQIVGCDHQLGSTVKEDNCGVCNGDGSTCRLVRGQYKSQLSATKS 201

Qy        244 EENVIAVPLGSRSVRITVKGPAHLFIESKTLQGSKGEHSFNSPGVFVVENTTVEFQRGSE 303
              :: |:|:|  |||  :|:  :|||  ||::|:||||||:|||:|  :|  |  |:|:|::|:||:   :
Db        202 DDTVVAIPYGSRHIRLVLKGPDHLYLETKTLQGTKGENSLSSTGTFLVDNSSVDFQKFPD 261

Qy        304 RQTFKIPGPLMADFIFKTRYTAAKDSVVQFFFYQPISHQWRQTDFFPCTVTCGGGYQLNS 363
              ::   ::  |||  ||||  |  |  :  :  ||  |||  |||||  |:||:|||||||:  ||||||||  |
Db        262 KETLRMAGPLTADFIVKIRNSGSADSTVQFIFYQPITHRWRETDFFPCSATCGGGYQLTS 321

Qy        364 AECVDLRKFEVVPEDHYCHYYPENVKPKPELKECSMDPCPSSDCFKEIMPYDHFQPLPRWE 423
              |||  |:|    |||  |  ||||||||:||||||:||::||||:                    |||
Db        322 AECYDLRSNRVVADQYCHYYPENIKPKPKLQECNLDPCPA-------------------RWE 364

Qy        424 HNPWTACSVSCGGGIQRRSFVCVEESMHGEILQVEEWKCMYAPKPKVMQTCNLFDCPKWI 483
                ||||||  |||||||  |:    ||||  :  |  :    ||||||||  ||    :  |  ||:||||||:
Db        365 ATPWTACSSSCGGGIQSRAVSCVEEDIQGHVISVEEWKCMYTPKMPIAQPCNIFDCPKWL 424

Qy        484 AMEWSQCTVTCGRGLRYRVVLCINHRGEHVGGCNPQLKLHIKEECVIPIPCYKPKEKSPV 543
              |  |||  |||||||:|||||||||||||:|||  |  ||||:|:  |  ||||||::|  ||||||||  ||
Db        425 AQEWSPCTVICGQGLRYRVVLCIDHRGMHTGGCSPKTKPHIKEECIVPIPCYKPKEKLPV 484

Qy        544 EAKLPWLKQAQELEETRIATEEPTFIPEPWSACSTTCGPGVQVREVKCRVLLTFTQTETE 603
              ||||||  ||||||||        :|||:|||:  ||||:  |||  |  |||  |:|:|||:|:|:   :
Db        485 EAKLPWFKQAQELEEGAAVSEEPSFIPKAWSACTVTCGVGTQVRIVRCQVLLSFSQSVAD 544

Qy        604 LPEEDCEGPKLPTERPCLLEACDESPASRELDIPLP         ---LPSLITYLWEYAGL 653
              ||  :||||||    ::|        ||      |  |  |  |:          |                :|  :    |||||  ||
Db        545 LPIDECEGPKPASQR-----ACYAGPCSGEIPEFNPDETDGLFGGLQDFDELYDWEYEGF 599

Qy        654 TECTATCLGGHQEAIAVCLHIQTQQTVNDSLCDMVHRPPAMSQACNTEPCPPRWHVGSWG 713
              |  |:  :|  ||  |||:    ||:  ||::        ::||          |||  :  ::||  :|||  ||  :|  |
Db        600 TKCSESCGGGVQEAVVSCLNKQTREPAFENLCVTSRRPPQLLKSCNLDPCPARWEIGKWS 659

Qy        714 PCSATCGVGIQTRDVYCLH-----PGETPAPPEQCSEEKPHALQACNQFDCPPGWHIEE 767
              |||  |||||:||||||:|  |             ||          :|  ||    ||    :||||:|:|||  |:    :
Db        660 PCSLTCGVGLQTRDVFCSHLLSREMNETVILADELCRQPKPSTVQACNRFNCPPAWYPAQ 719

Qy        768 WQQCSRTCGGGTQNRRVTCRQLLTDGSPLNLSDELCQGPKASSHKSCARTDCPPHLAVGD 827
              ||  ||||||||  |  |  |  |:|  :  |||||  |  :    |    |  :    ::|  :  |||       :  |
Db        720 WQPCSKTCGGGIVQRREVLCKLEMAIGSRLRLTLTDSASELACQQACRRDLPSLWLSS 779

Qy        828 WSKCSVSCGVGIQRRKQVCQRLAAKGRRIPLSEMMCRDLPGFPLVRSCQMPECSKIKSEM 887
              |::||  |||  |  |  |     :|:::       |         ::     :|     ||          :|  |  :    |::
Db        780 WIECSISCGEGIQIRSAICRKMLKIGLSIVVNSILCPPLPFSSSIRPCMLATCAR-PGRP 838

Qy        888 KTKLGEQGPQILSVQRVYIQTREEKRINLTIGSRAYLLPNISVIIKCPVRRFQKSLIQWE 947
               ||             |  |  :  ::||||||  :::::     :|     |||||  |:|::::||  ||  :|  ||  ||
Db        839 STK---HSPHIAAARKVYIQTRRQRKIHFVVGGFAYLLPKTAVVLRCPARRVRKPLIIWE 895

Qy        948 KDGRCLQNSKRLGITKSGSLKIHGLAAPDIGVYRCIAGSAQETVVLKLIGTDNRLIARP- 1006
              |||:  |  :|    :  :      |  ||||  |       |  |||  |  ||  |:|     |:||||  :  :|:|||
Db        896 EDGQHLISSTHVIVAPEGYLKIHELKPSDAGVYTCSAGSARLHEVIKLIGGNRKLVARPL 955

Qy       1007           --ALREPMREYPGMDHSEANSLGVTWHK  MRQMWNNKNDLYLDDDHISNQPF 1055
                           |  |:       :          |       |  :              |  :  |      |  ||  :|
Db        956 SPRSFEEVIACRKGGPKFALQTHFHQNCIFSNTSKAFKKSLAANPGSRY   EHLVSR--- 1010

Qy       1056 LRALLGHCSNSAGSTNSWELKNKQFEAAVKQ---GA---------YSMDTAQ--FDELIRN 1102
                   ||                 |||  ::                :        ||                  ::|  |  |      |:::  |
Db       1011 ---LLEQGGWPGELLASWEAQDSAERNTTSEEDPGAEQVLLHLPFTMVTEQRRLDDILGN 1067

Qy       1103 MSQLMETGEVSDDLASQLIYQLVAELAKAQPTHMQ--WRGIQEETPP           ---A 1148
              :||    |    |:  |     :      |:  ||    |:  ::        |             :      :    ||                       :
Db       1068 LSQQPF--FLRPDLYSKHLVAQLAQFIRPSHLEHQPTLLKPSERPISPVLLSPHKHVSGPS 1125

Qy       1149 AQLP----CPTGSVSQSSHAKNCCKLIFKPKCPVLMFL PQPPPISPNKTINSRIGNIVY 1203
               :  ||          |:  |    |:    |  |                              |  ::|:  |          :|  ::  :  :  :|  ||
Db       1126 SSLRTSSTGDAGGGSRRPHRK------------PTILRKISAAQQLSASEVV-THLGQTVA 1173

Qy       1204 ITKRTEVINILCDLITPSEATYTWIKDGTLLQPSVKIIMDGTGKTQTQNFTKKLQGIYTC 1263
               :       |      :   :   |:  |               |  :|   ::|      :|  |  :|:|                :||       |         :  |    |    |
Db       1174 LASGTLSVLLHCEAIGHPRPTISWARNGEEVQFSDRILLQPDDSLQILAPVEADVGFYTC 1233

Qy       1264 SVANHLGSDVESSSVLYAEAPVILSVERNITKFEHNHLSVVVGGIVEAALGANVTIRCPV 1323
               :     |  ||  |     |  :|       |  |::  :               :         |         ::|  :|         ::               |  ||||  |  |
Db       1234 NATNALGYDCVSIAVTLACKPLVKTTRMTVINTEKPAVTVDIGSTIKTVQGVNVTINCQV 1293

Qy       1324 KGVPQPNITWLKRGGSLSGNVSLLFNGSLLLQNVSLENEGTYVCIATNALGKAVATSVLH 1383
               |||:     :||  :             |  |:         |         |||||  |||        ::|   |    |    |    |    |:          ::  |
Db       1294 AGVPEAEVTWFRNKSKL-GSPHHLHEGSLLLTNVSSSDQGLYSCRAANLHGELTESTQLL 1352

Qy       1384 LLERRWPESRIVFLQGHKKYILQATNTRTNS------ -- --NDP----------TGEPP 1422
```

```
                :|        :::  ::         :  ||      |              |
Db       1353 ILDPPQVPIQLEDCFA     IILAAIGPNLPSVLISFDGIQLVLEPONSALLSPPIKGHEV 1408

Qy       1423 F                                                          --- 1423
              |
Db       1409 PNITWFHGGQPIVIATGITHHILAAGQILQVANLSGGSQGEFSCLAQNEAGVLMQKASLV 1468

Qy       1424  QEPPWEPGNWSHCSAPCGHGSAKIQRPQCVMANGQEVSEALCDHLQKPLAGFEPCNIRD 1482
               |: :|        : |::|::     |  :|:: |  |:: | |      :      |  ||
Db       1469 LQDYWWSVDRLATCSASCGNRGVQQPRLRCLLNSTEVNPAHCAGKVRPAVQPIACNRRD 1527

Qy       1483 CPARWFISVWSQCSVSCGEGYHSRQVTCKRTKANGTVQVVSPRACAPKDRPLGRKPCFG 1541
              ||:||   : || |: |||  |    :|:||||:: ||:|     ||    |      ||:  .
Db       1528 CPSRWMVPSWSACTRSCGGGVQTRRVPCQKLKASGTISTPVSNEMCTQVAKRPVDTQACNQ 1587

Qy       1542 HPCVQW  EPGNRCPGRCMGRAVRMQQRHTACQHNSS      DSNCDPRKRPTLRRNCTSG 1595
                  ||:|          :|  '  |:  :  :|  |    ||             |      |    : |  .
Db       1588 QDCVEWAPSSWGQCNGPCIGHHLAVQHRQVPCQTRDGTTLPSEQCSALPKPVSTQNCWSE 1647

Qy       1596 ACDVCWHPGPWRPCTAACGRGFQSRKVDCIHTRSCKPVAPQHCVQRPEEPSWRSCSLGPC 1653
              || |  |       |    |||  ||   ||||||:|:|:|  |:  |  |  |  :  |       :|  :|:  |  :  |
Db       1648 ACSVHWRVGDWSPCSATCGKGYQRRKVRCVHAKTNKAVPEHHCSWGPRPANWQRCNITP 1707

Qy       1654 SCDRPCTDCTHYCMEVKHLNCSCLEPYKQRPCQRP 1688
                 :  :|  ||'|  |:     |      '    :  :|  |:|  |  :|
Db       1708 CENMECRDTTRYCEKVKQCKTCQLSQFKSRCCQTCP 1742
```

```
RESULT 12
ABG65473
ID   ABG65473 standard; Protein; 1745 AA.
XX
AC   ABG65473;
XX
DT   27 AUG 2002  (first entry)
XX
DE   Human albumin fusion protein #2148.
XX
KW   Albumin fusion protein; therapeutic protein X; human albumin; HA;
KW   human serum albumin; HSA; cancer; reproductive disorder;
KW   digestive disorder; immune disorder; endocrine disorder;
KW   haematopoietic disorder; neural disorder; connective disorder;
KW   cytostatic; antiinfertility; antiinflammatory; antiulcer;
KW   immunomodulator; anti HIV; antidiabetic; haemostatic; nootropic;
KW   neuroprotective; antiparkinsonian; antimicrobial; neuroleptic;
KW   osteopathic; antiarthritic.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO200177137 A1.
XX
PD   18 OCT 2001.
XX
PF   12 APR 2001; 2001WO US11988.
XX
PR   12 APR 2000; 2000US 229358P.
PR   25 APR 2000; 2000US 199384P.
PR   21 DEC 2000; 2000US 256931P.
XX
PA   (HUMA ) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Haseltine WA;
XX
DR   WPI; 2002 010886/01.
XX
PT   New fusion protein for treating disease e.g. diabetes comprises an
PT   albumin fused to a therapeutic protein
XX
PS   Claim 1; Page 2042-2047; 210pp; English.
XX
CC   The present invention relates to albumin fusion proteins comprising a
CC   therapeutic protein X and human albumin (HA, also known as human serum
CC   albumin, HSA). The proteins are useful for treating a disease or
CC   disorder that may be modulated by therapeutic protein X. The albumin
CC   extends the shelf life of protein X, and may increase its biological
CC   in vitro/in vivo activity. The protein is useful for treating and
CC   diagnosing disorders such as cancer, reproductive disorders, digestive
CC   disorders (e.g. Crohn's disease, ulcerative colitis), immune disorders
CC   (e.g. acquired immunodeficiency syndrome, AIDS), endocrine disorders
CC   (e.g. diabetes), haematopoietic disorders, neural disorders
CC   (e.g. Alzheimer's, Parkinson's, Creutzfeldt Jacob disease,
CC   encephalomyelitis, meningitis, schizophrenia), and connective disorders
CC   (e.g. osteoporosis, arthritis). ABG63326-ABG65528 represent albumin
CC   fusion proteins of the invention.
XX
SQ   Sequence   1745 AA;

  Query Match             41.3%;  Score 8061;  DB 23;  Length 1745;
  Best Local Similarity   42.6%;  Pred. No. 1.6e-244;
  Matches  756;  Conservative  263;  Mismatches  552;  Indels  204;  Gaps   52;

Qy         63 LSKNIRSQAQKRGNWLAWQSQWSQCSRTCGGGASYSLRKCLSSKSCEGRNIRYRTCSNVDC 122
              :||   ||:||:||  ||||  ||:||||||||||||||||||:  :|||:||||:||||  ||
Db         22 GSKIARGEKGKGGWGAWGIWSQGSKTGGGASYSLRKCLSSKSCGGRNLFYRICSNVDC 81

Qy        124 PPDAEDPRAQGCSAYNEVQYQGHYYEWLPPYNDPAAPCALKCHAQGQNLVVELAPKVLDG 183
              ||:|  ||||||||||:|||:: |  ::||||    ||      |:::||  :      |  |  ||||||||
Db         82 PPEAGDPRAQQCSAHNDVKHHGQFYEWLPVSNDPDNPCSLKCQAKGTTLVVELAPKVLDG 141

Qy        184 TRCNTDSLDMCISGICQAVGCDRQLGSNAKEDNCGVCAGDGSTCRLVRGQSKSHVSPEKR 243
              |||  |:||||||||:|||  ||||  ||||    ||||||||  |||||||||||  ||  :|    |
Db        142 TRCYTESLDMCISGLCQIVGCDHQLGSTVKEDNCGVCNGDGSTCRLVRGQYKSQLSATKS 201

Qy        244 EENVIAVPLGSRSVRITVKGPAHLFIESKTLQGSKGEHSFNSPGVFVVENTTVEFQRGSE 303
              ::  |:|:|  ||:|  :  :  :  ',  ||::::|||:|::|  |:|  :       .:  :':::|:  |:  :
Db        202 DDTVVAIPYGSRHIRLVLKGPDHLYLETKTLQGTKGENSLSSTGTFLVDNSSVDFQKFPD 261

Qy        304 RQTFKIPGPLMADFIFKTRYTAAKDSVVQFFFYQPISHQWRQTDFFPCTVTCGGGYQLNS 363
              ::  ::  |||  ||||  |  |  :  :  ||  |||  |||||  |:||:||||||:  ||||||||  |
Db        262 KEILRMAGPLTADFIVKIRNSGSADSTVQFIFYQPIIHRWRETDFFPCSATCGGGYQLTS 321

Qy        364 AECVDIRLKRVVPDHYCHYYPENVKPKPKLKECSMDPCPSSDGFKEIMPYDHFQPLPRWE 423
              |||  |:|:  |  |  |  |||  |  |:|:||||:|:::||||:                   |||
Db        322 AECYDLRSNRVVADQYCHYYPENIKPKPKLQECNLDPCPA                   RWE 364

Qy        424 HNPWTACSVSCGGGIQRRSFVCVEESMHGEILQVEEWKCMYAPKPKVMQTCNLFDCPKWI 483
                 ||||||  ||||||  |:    ||||  :  |  :    ||||||||  ||    :  |  ||:||||||||:
Db        365 ATPWTACSSSCGGGIQSRAVSCVEEDIQGHVTSVEEWKCMYTPKMPIAQPCNIFDCPKWL 424

Qy        484 AMEWSQCTVTCGRGLRYRVVLCINHRGEHVGGCNPQLKLHIKEECVIPIPCYKPKEKSPV 543
              |  |||  ||||||:|||||  '  |||:'||  :  ||'::::  |  ,|,',|:::|  ||'  |:|||  ||
Db        425 AQEWSPCTVTCGQGLRYRVVLCIDHRGMHTGGCSPKTKPHIKEECIVPTPCYKPKEKLPV 484

Qy        544 EAKLPWLKQAQELEETRIATEEPTFIPEPWSACSTTCGPGVQVREVKCRVLLTFTQTETE 603
              ||||||  ||||||||        :  ||:|||:  |||':  |:|  |  |||  |:|:||·:|:|:  :
Db        485 EAKLPWFKQAQELEEGAAVSEEPSFIPEAWSACTVTCGVGTQVRIVRCQVLLSFSQSVAD 544

Qy        604 LPEEECEGPKLPTERPCLLEACDESPASRELDIPLP          EDCETDYDWEYAGF 653
              ||  :|||||||  ::|        |    |  |  |:      |           :|  :  |||||  ||
Db        545 LPIDECEGPKPASQR       ACYAGPCSGEIPEFNPDETDGLFGGLQDFDELYDWEYEGF 599

Qy        654 TPCTATCLGGHQEATAVCLHIQTQQTVNDSLCDMVHRPPAMSQACNTEPCPPRWHVGSWG 713
              |  |:  :|  ||  |||:    ||:  ||::       ::||       |||  :  ::||  :|||  ||  :|  |
Db        600 TKCSESCGGGVQEAVVSCLNKQTREPAEENLCVTSRRPPQLLKSCNLDPCPARWEIGKWS 659

Qy        714 PCSATCGVGIQTRDVYCLH         PGETPAPPEECRDEKPHALQACNQFDCPPGWHIEE 767
              |||  ||||!:|||||:|  |             ||        :|  |'      |'     :||!|:|:|||  |:    :
Db        660 PCSLTCGVGLQTRDVFCSHLLSREMNETVILADELCRQPKPSTVQACNRFNCPPAWYPAQ 719

Qy        768 WQQCSRTCGGGTQNRRVTCRQLLTDGSFLNLSDELCQGPKASSHKSCARTDCPPHLAVGD 827
              ||  ||||||||  |  |  |  |:'  :  ||||!  |  :    |      |  :    ::'  :  ||'      :  |
Db        720 WQPCSRTCGGGVQKRLVLCKQRMADGSFLELPETFCSASKPACQQACKKDDCPSEWLLSD 779

Qy        828 WSKCSVSCGVGIQRRKQVCQRLAAKGRRIPLSEMMCRDLPGLPLVRSCQMPECSKIKSEM 887
              |::||  |,|  |  !  |    :!:::      |         ::      .·   |·         :|  |  :   ::·
```

```
Db        780 WTECSISCGEGIQTRSAICKKMLKTGLSTVVNSTLCPPLPFSSSIRPCMLATCAR-PGRP 838

Qy        888 KTKICEQCPQILSVQRVYIQTPEEKRINLTIGSPAYLLPNTSVIIKCPVRRPQKSLIQWE 947
                ||      | |  :  ::||||||  :::::   :|   ||||| |:|:::|| ||  :| || ||
Db        839 STK---HSPHIAAARKVYIQTPRQPKLHFVVGGFAYLLPKTAVVLRCPARRVRKPLITWE 895

Qy        948 KDGRCLQNSKRLGIIKSGSLKIHGLAAPDGVYRCIAGSAQEIVVLKLIGTDNRLIARP- 1006
              |||:  | :|   :  :    |  ||||  |    |  |||  |  ||  |:|   |:||||  : :|:|||
Db        896 KDGQHLISSTHVTVAPFGYLKIHRLKPSDAGVYTCSAGPAREHFVIKLIGGNRKLVARPL 955

Qy       1007 --------ALREPMREYPGMDHSEANSLGVTWHK  MPQMWNNKNTTYLLDPHIPNQPF 1055
                        | |:   :      |    | :      |    | :   |      |   || :|
Db        956 SPRSEEEVLAGRKGGPKEALQTHKHQNGIFSNGSKAEKRGLAANPGSRY--DDLVSR--- 1010

Qy       1056 LRALLGHCSNSAGSTNSWELKNKQFEAAVKQ---GA--------YSMDTAQ--FDELIRN 1102
                   ||              ||| ::          :   ||          ::| |  |     |::: |
Db       1011 ---LLEQGGWPGELLASWEAQDSAERNTTSEEDPGAEQVLLHLPFTMVTEQRRLDDILGN 1067

Qy       1103 MSQLMEIGEVSDDLASQLIYQLVAPLAKAQPTHMQ  WPGIQEETPP-------------A 1148
              :||  |   |:  |   :   |:  ||   |: ::   |        :   :  | |              :
Db       1068 LSQQPE--FLRDLYSKHLVAQLAQEIFRSHLEHQDILLKPSERRTSPVTLSPHKHVSGFS 1125

Qy       1149 AQIP      GPTGGVSLSSHAKNCGKITEEPKGPVIMKL  SQPPSISFNKTTNSPIGNIVY 1202
              : ||      |:  :  |:   | |                  | ::|:  |      :|  ::  :  :  :| ||
Db       1126 SSLRTSSTGDAGGGSRRPHRK------     PTILRKISAAQQLSASEVV THLGQTVA 1173

Qy       1204 ITKPTEVINIIGCLITPSPATYTWTKPGTLLQPSVKIIIEGTGKIQIQNFTKKLQGIYEC 1263
              :    |   :  :  |.  |          | :|  ::|    :| |  :|:|        :||   |     :  |  |  |
Db       1174 LASGTLSVLLHCEAIGHPRPTISWARNGEEVQFSDRILLQPDDSLQILAPVEADVGFYTC 1233

Qy       1264 SVANHLGSDVESSSVLYAEAPVILSVERNITKPEHNHLSVVVGGIVEAALGANVTIRCPV 1323
              :   |  ||  |    |  :|    |    |::  :        :      |      ::|  :|      ::      | ||||  |  |
Db       1234 NATNALGYDSVSIAVTLAGKPLVKTSRMTVINTEKPAVTVDIGSTIKTVQGVNVTINCQV 1293

Qy       1324 KGVPQPNITWLKRGGSLSGNVSLLFNGSLLLQNVSLENEGTYVCIATNALGKAVATSVLH 1383
                 |||:   :||  :           |  ::        |      ||||| |||     ::|  |  |  |  |    |:          ::  |
Db       1294 AGVPEAEVTWFRNKSKLGSPHHLHEGSLLLTNVSSSDQGLYSCRAANLHGELTESTQLL 1352

Qy       1384 IIERRWPESRIVPLQGHKKYILQATNTPTNS------------NDP----     TGEPP 1422
              :|:          :::    ::           :| ||          |                ||                ||
Db       1353 ILDPPQVPTQLEDIRA----LLAATGPNLPSVLTSPLGTQLVLDPGNSALLGCPIKGHPV 1408

Qy       1423 P------------------------------------------------------------ 1423
              |
Db       1409 PNITWFHGGQPIVTATGLTHHILAAGQILQVANLSGGSQGEFSCLAQNEAGVLMQKASLV 1468

Qy       1424 -QEPTWEPGNWSHCSATCGHLGARIQRPQCVMANGQEVSEAICDHLQKPLAGFEPCNIRD 1482
                |:  :|           :  ||:|||:  |  :     |  :|::  |     ||:  |  |          :|            ||  ||
Db       1469 IQDYWWSVDRLATCSASCGNRGVQQPRLRCLL NSTEVNPAHCAGKVRPAVQPIACNRRD 1527

Qy       1483 CPARWFTSVWSQCSVSCGEGYHSRQVTCKRTKANGTVQVVSPRACAP-KDRPLGRKPCPG 1541
              ||:||     :  ||  |:  |||  |     :|:|||::  ||:|         ||      |          ||:    :  |
Db       1528 CPSRWMVTSWSACTRSCGGGVQTRRVTCQKLKASGISTPVSNDMCTQVAKRPVDTQACNQ 1587

Qy       1542 HPCVQW--EPGNRCPGRCMGRAVRMQQRHTACQHNSS----DSNCDDRKRPTLRRNCTSG 1595
                 ||:|           :|  |  |:|    :  :|  |     ||                       |       ||        :||  |
Db       1588 QLCVEWAFSSWGQCNGPCIGPHLAVQHRQVFCQTRDGITLPSEQCSALPRPVSTQNCWSE 1647

Qy       1596 ACDVCWHTGPWKPCTAACGR-GFQSRKVDCIHTRSCKPVAKRHCVQKKKPISWRHC-LGP 1653
              || |  |         |   ||: ||    |||||:|:|:| |  |:  |  |  .     |          .| .|.  |  :  |
Db       1648 ACSVHWRVSLWTLCTATCGNYGFQSRRVECVHARTNKAVPEHLCSWGPRPANWQRCNITP 1707

Qy       1654 SCDRDCTDTTHYCMFVKHLNLCSLDRYKQRCCQSC 1688
                  :  :|  |||  ||    ||  |  ||  |  ::|  |||  :|
Db       1708 CENMECRDTTRYCEKVKQLKLCQLSQFKSRCCGTC 1742


RESULT 13
ABB10246
ID   ABB10246 standard; Protein; 1766 AA.
XX
AC   ABB10246;
XX
DT   10-JAN-2002  (first entry)
XX
DE   Human cDNA SEQ ID NO: 554.
XX
KW   Human, gene therapy, neural disorder, immune system disorder;
KW   muscular disorder; reproductive disorder; gastrointestinal disorder;
KW   pulmonary disorder; cardiovascular disorder; renal disorder;
KW   proliferative disorder; inflammation.
XX
OS   Homo sapiens.
XX
PN   WO200154474-A2.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01349.
XX
PR   31-JAN-2000; 2000US-179065P.
PR   04-FEB-2000; 2000US-180628P.
PR   24-FEB-2000; 2000US-184664P.
PR   02-MAR-2000; 2000US-186350P.
PR   16-MAR-2000; 2000US-189874P.
PR   17-MAR-2000; 2000US-190076P.
PR   18-APR-2000; 2000US-198123P.
PR   19-MAY-2000; 2000US-205515P.
PR   07-JUN-2000; 2000US-209467P.
PR   28-JUN-2000; 2000US-211886P.
PR   30-JUN-2000; 2000US-215135P.
PR   07-JUL-2000; 2000US-216647P.
PR   07-JUL-2000; 2000US-216880P.
PR   11-JUL-2000; 2000US-217487P.
PR   11-JUL-2000; 2000US-217496P.
PR   14-JUL-2000; 2000US-218290P.
PR   26-JUL-2000; 2000US-220963P.
PR   26-JUL-2000; 2000US-220964P.
PR   14-AUG-2000; 2000US-224518P.
PR   14-AUG-2000; 2000US-224519P.
PR   14-AUG-2000; 2000US-225213P.
PR   14-AUG-2000; 2000US-225214P.
PR   14-AUG-2000; 2000US-225266P.
PR   14-AUG-2000; 2000US-225267P.
PR   14-AUG-2000; 2000US-225268P.
PR   14-AUG-2000; 2000US-225270P.
PR   14-AUG-2000; 2000US-225447P.
PR   14-AUG-2000; 2000US-225757P.
PR   14-AUG-2000; 2000US-225758P.
PR   14-AUG-2000; 2000US-225759P.
PR   18-AUG-2000; 2000US-226279P.
PR   22-AUG-2000; 2000US-226681P.
PR   22-AUG-2000; 2000US-226868P.
PR   22-AUG-2000; 2000US-227182P.
PR   23-AUG-2000; 2000US-227009P.
PR   30-AUG-2000; 2000US-228924P.
PR   01-SEP-2000; 2000US-229287P.
PR   01-SEP-2000; 2000US-229343P.
PR   01-SEP-2000; 2000US-229344P.
PR   01-SEP-2000; 2000US-229345P.
PR   05-SEP-2000; 2000US-229509P.
PR   05-SEP-2000; 2000US-229513P.
PR   06-SEP-2000; 2000US-230437P.
PR   06-SEP-2000; 2000US-230438P.
PR   08-SEP-2000; 2000US-231242P.
PR   08-SEP-2000; 2000US-231243P.
PR   08-SEP-2000; 2000US-231244P.
PR   08-SEP-2000; 2000US-231413P.
PR   08-SEP-2000; 2000US-231414P.
PR   08-SEP-2000; 2000US-232080P.
PR   08-SEP-2000; 2000US-232081P.
PR   12-SEP-2000; 2000US-231968P.
PR   14-SEP-2000; 2000US-232397P.
PR   14-SEP-2000; 2000US-232398P.
PR   14-SEP-2000; 2000US-232399P.
PR   14-SEP-2000; 2000US-232400P.
PR   14 SEP 2000; 2000US 232471P.
PR   14 SEP 2000; 2000US 233063P.
PR   14 SEP 2000; 2000US 233064P.
PR   14 SEP 2000; 2000US 233065P.
PR   21 SEP 2000; 2000US 234223P.
PR   21 SEP 2000; 2000US 234274P.
PR   25 SEP 2000; 2000US 234997P.
PR   25 SEP 2000; 2000US 234998P.
PR   26 SEP 2000; 2000US 235484P.
PR   27 SEP 2000; 2000US 235834P.
PR   27 SEP 2000; 2000US 235836P.
PR   29 SEP 2000; 2000US 236327P.
PR   29 SEP 2000; 2000US 236367P.
PR   29 SEP 2000; 2000US 236368P.
PR   29 SEP 2000; 2000US 236369P.
PR   29 SEP 2000; 2000US 236370P.
PR   02 OCT 2000; 2000US 236802P.
PR   02 OCT 2000; 2000US 237037P.
PR   02 OCT 2000; 2000US 237038P.
PR   02 OCT 2000; 2000US 237039P.
PR   02 OCT 2000; 2000US 237040P.
PR   13 OCT 2000; 2000US 239935P.
PR   13 OCT 2000; 2000US 239937P.
PR   20 OCT 2000; 2000US 240960P.
PR   20 OCT 2000; 2000US 241221P.
PR   20 OCT 2000; 2000US 241785P.
PR   20 OCT 2000; 2000US 241786P.
PR   20 OCT 2000; 2000US 241787P.
PR   20 OCT 2000; 2000US 241808P.
PR   20 OCT 2000; 2000US 241809P.
PR   20 OCT 2000; 2000US 241826P.
PR   01 NOV 2000; 2000US 244617P.
PR   08 NOV 2000; 2000US 246474P.
PR   08 NOV 2000; 2000US 246475P.
PR   08 NOV 2000; 2000US 246476P.
PR   08 NOV 2000; 2000US 246477P.
PR   08 NOV 2000; 2000US 246478P.
PR   08 NOV 2000; 2000US 246523P.
PR   08 NOV 2000; 2000US 246524P.
PR   08 NOV 2000; 2000US 246525P.
PR   08 NOV 2000; 2000US 246526P.
PR   08 NOV 2000; 2000US 246527P.
PR   08 NOV 2000; 2000US 246528P.
PR   08 NOV 2000; 2000US 246532P.
PR   08 NOV 2000; 2000US 246609P.
PR   08 NOV 2000; 2000US 246610P.
PR   08 NOV 2000; 2000US 246611P.
PR   08 NOV 2000; 2000US 246613P.
PR   17 NOV 2000; 2000US 249207P.
PR   17 NOV 2000; 2000US 249208P.
PR   17 NOV 2000; 2000US 249209P.
PR   17 NOV 2000; 2000US 249210P.
PR   17 NOV 2000; 2000US 249211P.
PR   17 NOV 2000; 2000US 249212P.
PR   17 NOV 2000; 2000US 249213P.
PR   17 NOV 2000; 2000US 249214P.
PR   17 NOV 2000; 2000US 249215P.
PR   17 NOV 2000; 2000US 249216P.
PR   17 NOV 2000; 2000US 249217P.
PR   17 NOV 2000; 2000US 249218P.
PR   17 NOV 2000; 2000US 249244P.
PR   17 NOV 2000; 2000US 249245P.
PR   17 NOV 2000; 2000US 249264P.
PR   17 NOV 2000; 2000US 249265P.
PR   17 NOV 2000; 2000US 249297P.
PR   17 NOV 2000; 2000US 249299P.
PR   17 NOV 2000; 2000US 249300P.
PR   01 DEC 2000; 2000US 250160P.
PR   01 DEC 2000; 2000US 250391P.
PR   05 DEC 2000; 2000US 251030P.
PR   05 DEC 2000; 2000US 251868P.
PR   05 DEC 2000; 2000US 256714P.
PR   06 DEC 2000; 2000US 251479P.
PR   08 DEC 2000; 2000US 251856P.
PR   08 DEC 2000; 2000US 251868P.
PR   08 DEC 2000; 2000US 251869P.
PR   08 DEC 2000; 2000US 251989P.
PR   08 DEC 2000; 2000US 251990P.
PR   11 DEC 2000; 2000US 254097P.
PR   05 JAN 2001; 2001US 259678P.
XX
PA   (HUMA ) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Barash SC,  Ruben SM;
XX
DR   WPI; 2001-476161/51.
DR   N-PSDB; ABA06468.
XX
PT   Isolated nucleic acid molecule encoding an inflammation-associated
PT   polypeptide is used in preventing, treating or ameliorating a medical
PT   condition
XX
PS   Claim 11; SEQ ID NO: 554; 85pp + Sequence Listing; English.
XX
CC   The present invention provides human cDNAs, proteins and related genomic
CC   DNAs. These can be used in the treatment of neural, immune system,
CC   muscular, reproductive, gastrointestinal, pulmonary, cardiovascular
CC   renal and proliferative disorders and inflammation. The present sequence
CC   is a protein of the invention.
XX
SQ   Sequence   1766 AA;

  Query Match             41.3%;  Score 3860.5;  DB 22;  Length 1766;
  Best Local Similarity   42.1%;  Pred. No. 1.7e-244;
  Matches  765;  Conservative  266;  Mismatches  579;  Indels  207;  Gaps   33;

Qy        25 IAEKSPGAYFLPEFALSPQGSF    LEDTTGEQFLLYRYDDQTSRNIRSDEDKIQNWDAW 81
             |||:    || : :                 | |    |    :     :||   ||:| :|| ||||
Db         1 IAEEHLAAYSVSDSDSGKDPSMECCRRATPGLLLLFLAFLLLSSRIARSEEDRDGLWDAW 60

Qy        82 GDWSDCSRTCGGGASYSLRRCLTGRNCEGQNIRYKTCSNHDCPPDAEDFRAQQCSAYNDV 141
             | ||:|||||||||||||||| |: ::|||:||||:||||  |||||:| ||||  |||||:||
Db        61 CPWSECSRTCGGGASYSLRRCLSSKSCEGRNIRYRTCSNVDCPPEAGDFRAQQCSAHNDV 120

Qy       142 QYQGHYYEWLPRYNDPAAPCALKCHAQGQNLVVELAPKVLDGTRCNTDSLDMCISGICQA 201
             :: | :|||||   |||  || |:| |:|  :|||||||||:||||| |:|||  |||:||
Db       121 KHHGQFYEWLPVSNDPDNPCSLKCQAKGTTLVVELAPKVLDGTRCYTESLDMCISGLCQI 180

Qy       202 VGCDRQLGSNAKEDNCGVCAGDGSTCRLVRGQSKSHVSPEKREENVIAVPLGSRSVRITV 261
             |||| ||||   ||||||| ||||||||||||| || :|   | :: |:|:| :|| :|: :
Db       181 VGCDHQLGSTVKEDNCGVCNGDGSTCRLVRGQYKSQLSAIKSDDTVVAIPYGSRHIRLVL 240

Qy       262 KGPAHLFIESKTLQGSKGEHSFNSPGVFVVENTTVEFQRGSERQTFKIPGPLMADFIFKT 321
             ||| ||::|:|||||:||| :  :|  | |:|:|:|:|  :   :::  :: ||  |||| |
Db       241 KGPDHLYLETKTLQGTKGENSLSSTGTFLVDNSSVDFQKFPDKEILRMAGPLTADFIVKI 300

Qy       322 RYTAAKDSVVQFFFYQPISHQWRQTDFFPCTVTCGGGYQLNSAECVDIRLKRVVPDHYCH 381
             | : : || |:| ||||| |:||:||||||: |||||||| |||| |:|  :|| | |||
Db       301 RNSGSADSTVQFIFYQPIIHRWRETDFFPCSATCGGGYQLTSAECYDLRSNRVVADQYCH 360

Qy       382 YYPENVKPKPKLKECSMDPCPSSDGFKEIMPYDHFQPLPRWEHNPWTACSVSCGGGIQRR 441
             ||||:|| | |:||::||||:                    |:|  ||:|||  |||:|||| |
Db       361 YYPENIKPKPKLQECNLDPCPA                    PWLATPWTACSSSCGGGIQSR 403

Qy       442 SFVCVEESMHGEIIQVEEWKCMYAPKPKVMQTCNLFDCPKWIAMEWSQCTVTCGRGLRYR 501
             :  |||| : | :  ||||||||  ||  : |  ||:||||| :| |||  |||:|||||
Db       404 AVSCVEEDIQGHVTSVEEWKCMYTPKMPIAQPCNIFDCPKWLAQEWSPCTVTCGQGLRYR 463

Qy       502 VVLCINHRGEHVGGCNPQLKLHIKEECVIPIPCYKPKEKSPVEAKLPWLKQAQELEETRI 561
             |||||:||| | |||:|: | ||||||:|||||||||||  ||||||||:||||||| 
Db       464 VVLCIDHRGMHTGGCSPKTKPHIKEECIVPIPCYKPKEKLPVEAKLPWFKQAQELEEGAA 523

Qy       562 ATEEPTFIPEPWSACSTTCGPGVQVREVKCRVLLTFTQTETELPEEECEGPKLPTERPCL 621
              :|||:|||  :|:|: |:  | |||:|:|:|||:|:|:  :: :|||  :|||||| ::: 
Db       524 VSEEPSFIPXAWSACTVTCGVGTQVRIVRCQVLLSFSQSVADLPIDECEGPKPASQR    580
```

```
Qy        622 LEACDESPASRELGIPLP-----------EDSETTYDWEYAGFTPCTATCLGGHQEAIAVC 671
                ||     | | | :        |               :|  :   ||||| |||  |:  :| || |||:   |
Db        581 --ACYAGPCSGEIPEFNPDEIDGLFGGLQDFDELYDWEYEGFTKCSESCGGGVQEAVVSC 638

Qy        672 LHIQTQQIVNDSLCDMVHRPPAMSQACNIEPCPPRWHVGSWGPCSAICGVGIQTRDVYCL 731
              |: ||::      ::||       ||| :  ::|| :||| || :| | ||| |||||:|||||:|
Db        639 LNKQTRELALENGSVISKRPFQLLKSCNLDPCPARWEIGKWSPCSLICGVGLQTRDVFCS 698

Qy        732 H-      PGETPAPPEE-CRDEKPHALQACNQFDCPPGWHIEEWQQCSRTCGGGTQNRRVT 785
              |        ||      :| ||    ||    :||||:|:||| |:    :|| ||||||||  | | |
Db        699 HLLSREMNETVILADELCRQPKPSTVQACNRFNCPPAWYPAQWQPCSRTCGGGVQKREVL 758

Qy        786 CRQLLTDGSFLNLSDELCQGPKASSHKSCARTDCPPHLAVGDWSKCSVSCGVGIQRRKQV 845
              |:|  : ||||| |  :    !      |  :    ::| : |||           : ||::|| ||| | | |   :
Db        759 CKQRMADGSPLPLPETECSAFPPACQQACRKEDCPLWLLSDWIECSISCGEGTQTRSAI 818

Qy        846 CQRLAAKGRRIPLSEMMCRDLPGFPLVRSCQMPECSKIKSEMKTKLGEQGPQILSVQRVY 905
              |:::      |        ::    :|   ||          :|  |  :  |::          ||          | |  : ::||
Db        819 CRKMLKICLSTVVNSTLCPPLPPSSCIRPCMLATCAR IGKLSIK    HSPHIAAARKVY 874

Qy        906 IQTPEEKRINLTIGSRAYLLFNTSVIIKCPVRRFQKSLIQWEKDGRCLQNSKRLGITKSG 965
              ||||  :::::     :|    |||||  |:|:::|| ||  :|  || ||||:  |  :|    :  :     |
Db        875 IQTRRQRKLHFVVGGFAYLLPKTAVVLRCPARRVRKPLITWEKDGQHLISSIHVTVAPFG 934

Qy        966 SLKIHGLAAPDIGVYRCIAGSAQETVVLKLIGTDNRLIARP-----------ALREPMREY 1015
               |||  |         | ||  |  ||  |:|      |:||||  :  :|:|||                 | |:     :
Db        935 YLKIHPLRPSDAGVYTCSAGPAREHFVIKLIGGNRKLVARPLSPRSEEEVLAGRKGGPKE 994

Qy       1016 PGMDHSEANSLGVTWHK--MRQMWNNKNDLYLDDDHISNQPFLRALLGHCSNSAGSTNSW 1073
                        |      |  :         |     |  :     |           |    || :|                    ||                    ||
Db        995 ALQTHKHQNGIFSNGSKAEKRGLAANPCSRY   DDLVSK         LLEQGGWPGELLASW 1046

Qy       1074 ELKNRQIEAAVKL       GA        YEMDTAQ  FDLLIRNMSQLMEIGEVSDDLASQL 1120
               |  : :             :      ||               : :|  |  |         |: : :   |:||     |     |:    |         :    |
Db       1047 EAQDSAERNTTSEEDPGAEQVLLHLPFTMVTEQRRLDDILGNLSQQPF--ELRDLYSKHL 1104

Qy       1121 IYQLVAELAKAQPIHMQ--WRCIQEETPP-------   ---AAQLR----GETGSVSQSS 1162
               :  ||    |: ::       |         :    :    | |                          :: ||          |:  |    |:
Db       1105 VAQIAQEIPRSHIPHQPTIIKPSEPRTSPVTLSPHKHVSGPSSSLKISSIGDAGGGSRRP 1164

Qy       1163 HAKNSCKLIPKPKGPVLMPQ-SQPPSISENKTINPKIGNIVYIIPRIEVINILCDLIIPS 1221
               | |                   | ::|: |        :|  ::  :   :  :|  ||  :        |     :   :  |:  |
Db       1165 HRK------------PTILRKISAAQQLSASEVV THLGQTVALASGTLSVLLHCEAIGHP 1212

Qy       1222 EATYTWTKDGTLLQPSVKIILDGTGKIQIQNPTRKEQGIYECSVANHLGSDVESSSVLYA 1281
                 |  :|  ::|     :| |  :|:|          :||   |           :  |  |  |:      ||| |     |  :|    |
Db       1213 RPTISWARNGEEVQFSDRILLQPDDSLQILAPVEADVGFYTCNATNALGYDSVSIAVTLA 1272

Qy       1282 EAPVILSVERNIIKPEHNHLSVVVGGIVEAALLANVIIRCPVKGVPQPNITWLKRGGSLS 1341
                 |:::  :        :      |          ::|  :|    ::          |  ||||  |   |   |||:      :||  :             |
Db       1273 GKPLVKTSRMTVINTEKPAVTVDIGSTIKTVQGVNVTINCQVAGVPEAEVIWFRNKSKL- 1331

Qy       1342 GNVSLLFNGSLLLQNVSLENEGIYVCIAINALGKAVAISVLHLLERRWPESRIVFLQGHK 1401
              |:        |      ||||: |||    ::|   |   |   |   |   |    |:          ::   |   :|:                :::    ::
Db       1332 GSPHHLHEGSLLLTNVSSSDQGLYSCRAANLHGELTESTQLLILDPPQVPTQLEDIRA-  1389

Qy       1402 KTILQATNIKINS          --NDP-------   IGEPPP------------------ 1423
                  :|  ||          |                      ||                      |  |  |
Db       1390   LLAATGPNLPSVLTSPLGTQLVLDPGNSALLGCPIKGHPVPNITWFHGGQPIVTATGL 1447

Qy       1424 ------------------------------------------QEPPWEPGNWSHCSATC 1440
                                                          |:  :|              :   |||:|
Db       1448 THHILAAGQILQVANLSGGSQGEFSCLAQNEAGVLMQKASLVIQDYWWSVDRLATCSASC 1507

Qy       1441 GHLGARIQRPQCVMANGQEVSEALCDHLQKPLAGFEPCNIRDCPARWFTSVWSQCSVSCG 1500
              |:  |   :     |  :|::  |     ||:  |  |             :|                  ||  ||||:||    :  ||  |:  |||
Db       1508 GNRGVQQPRLRCLL-NSTEVNPAHCAGKVRPAVQPIACNRRDCPSRWMVTSWSACTRSCG 1566

Qy       1501 EGYHSRQVICKRIKANGTVQVVSPRACAP-KDRPLGRKPCFGHPCVQW--EPGNRCPGRC 1557
                |       :|:|||:: ||:|           ||       |               ||:     :   |               |||:|                  :|  |   |
Db       1567 GGVQTRRVTCQKLKASGISIPVSNDMCTQVAKRPVDTQACNQQLCVEWAFSSWGQCNGPC 1626

Qy       1558 MGRAVRMQQRHTACQHNSS----DSNCDDRKRPTLRRNCISGACDVCWHTGPWKPCTAAC 1613
```

```
                    :|   :  :|  |      ||                    |        ||       :||  |  ||  |  |       |    ||| |
Db       1627 IGPHLAVQHPQVPCQTRDCITLPSEQCSALPRPVCILQNCWSEACSVHWKVSLWTLCTATC 1686

Qy       1614 GP GPQSPKVPCIHTPSCKPVAKSHCVLKKKPISWRHC LGPSCDEPCTETTHYCMLVKH 1671
                |    ||||:|:|:|    |:   |  |  :    |           :|  :|:   |  :   |     :   :|   |||  ||     ||
Db       1687 CNYCFQSPRVPCVHAPTNKAVPPHLCCWCLRPANWQRCNIIPCPNMPCRCTTKYCIFVKQ 1746

Qy       1672 LNLCSLDRYKQRCCQSC 1688
                | ||  |  ::|  |||  :|
Db       1747 LKLCQLSQFKSRCCGTC 1763


RESULT 14
AAE07867
ID   AAE07867 standard; Protein; 650 AA.
XX
AC   AAE07867;
XX
DT   01-NOV-2001  (first entry)
XX
DE   Novel human protein (NHP) #5.
XX
KW   Novel human protein; NHP; thrombospondin; gene therapy; cancer;
KW   cytostatic; antisense therapy; angiogenesis; biological disorder.
XX
OS   Homo sapiens.
XX
PN   WO200161011-A2.
XX
PD   23-AUG-2001.
XX
PF   15-FEB-2001; 2001WO-US05290.
XX
PR   17-FEB-2000; 2000US-0183282.
XX
PA   (LEXI-) LEXICON GENETICS INC.
XX
PI   Donoho G,  Scoville J,  Turner CA,  Friedrich G,  Zambrowicz B;
PI   Sands AT;
XX
DR   WPI; 2001-514776/56.
DR   N-PSDB; AAD14368.
XX
PT   An isolated nucleic acid encoding a thrombospondin useful as a
PT   hybridization probe and gene therapy treatments of cancer -
XX
PS   Disclosure; Page 41-42; 56pp; English.
XX
CC   The invention relates to novel human polynucleotides encoding proteins
CC   that share sequence similarity with animal proteins having thrombospondin
CC   repeats. NHPs are expressed in human cell lines, pituitary, lymph nodes,
CC   prostrate, testis, adrenal gland, uterus, foetal kidney, foetal lung and
CC   gene trapped human cells. NHPs are thrombospondins useful for treating
CC   biological disorders involving angiogenesis, cancer and development and
CC   also in pharmacogenomic applications. NHPs are useful as a hybridisation
CC   probe for screening libraries, assessing gene expression patterns and
CC   also in gene therapy. Proteins having thrombospondin repeats act as
CC   receptors, secreted extracellular matrix proteins and proteases. The
CC   present sequence is novel human protein (NHP).
XX
SQ   Sequence   650 AA;

  Query Match             37.4%;  Score 3492;  DB 22;  Length 650;
  Best Local Similarity   89.5%;  Pred. No. 6.2e-221;
  Matches  632;  Conservative    0;  Mismatches    0;  Indels   74;  Gaps    1;

Qy          1 MASWTSPWWVLIGMVFMHSPLPQTTAEKSPGAYFLPEFALSPQGSFLEDTTGEQFLTYRY 60
              |||||||||||||||||||||||||||||
Db          1 MASWTSPWWVLIGMVFMHSPLPQTTAEKSP------------------------------ 30

Qy         61 DPQTSRNTRSPDPKPCNWPAWGEWSLCSPTCGLGTACTIRRSLIGRNCEGQNIRYKICSN 120
                                                            ||||||||||||||||
Db         31 --------------------------------------------GRNCEGQNIRYKICSN 46
```

```
Qy        121 HDCPPDAEDFRAQQCSAYNDVQYQGHYYEWLPFRNDPAAPCALKCHAQGQNLVVELAPKV 180
              ||:|||||||||||:||||   :|||:||||||||||||| ||:| || | |:|:: |||:  |:
Db         47 HDCPPDAEDFRAQQCSAYNDVQYQGHYYEWLPFYNDPAAPCALKCHAQGQNLVVELAPKV 106

Qy        181 LDGIRCNIDSLDMCISGICQAVGCDRQLGSNAKEDNCGVCAGDGSTCRLVRGQSKSHVSP 240
              ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        107 LDGIRCNIDSLDMCISGICQAVGCDRQLGSNAKEDNCGVCAGDGSTCRLVRGQSKSHVSP 166

Qy        241 EKREENVIAVPLGSRSVRIIVKGPAHLFIESKTLQGSKGEHSFNSPGVFVVENTIVEFQR 300
              ||||||||||||||||||||.||||||||||||||||||||||||||||||||| ||||||
Db        167 EKREENVIAVPLGSRSVRIIVKGPAHLFIESKTLQGSKGEHCFNSPGVFVVENTIVEFQR 226

Qy        301 GSERQTFKIPGPLMADFIFKTRYTAAKDSVVQFFFYQPISHQWRQTDFFPCIVTCGGGYQ 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 GSERQTFKIPGPLMADFIFKTRYTAAKDSVVQFFFYQPISHQWRQTDFFPCIVTCGGGYQ 286

Qy        361 LNSAECVDIRLKRVVPDHYCHYYPENVKPKPKLKECSMDPCPSSDGFKEIMPYDHFQPLP 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 LNSAECVDIRLKRVVPDHYCHYYPENVKPKPKLKECSMDPCPSSDGFKEIMPYDHFQPLP 346

Qy        421 RWEHNPWIACSVSSCGGIQRRSFVCVEESMHGEILQVEEWKCMYAPKPKVMQTCNLFDCP 480
              ||| |||||||  :: | |  :::|||||||||:||||||||:||||||| : :|  :|:|
Db        347 RWEHNPWIACSVSSCGGIQRRSFVCVEESMHGEILQVEEWKCMYAPKPKVMQTCNLFDCP 406

Qy        481 KWIAMEWSQCTVTCGRGLRYRVVLCINHRGEHVGGCNPQLKLHIKEECVIPIPCYKPKEK 540
              ||| |||||||:| |||| |||||||||||||||||||||||||||||||||||||:|:||
Db        407 KWIAMEWSQCTVTCGRGLRYRVVLCINHRGEHVGGCNPQLKLHIKEECVIPIPCYKPKEK 466

Qy        541 SPVEAKLPWLKQAQELEETRIATEEPTFIPEPWSACSTTCGPGVQVREVKCRVLLTFTQT 600
              |||:|||||||.| |||: :.|.|||||||||||||||| ||||||||:: :  |::|.||
Db        467 SPVEAKLPWLKQAQELEETRIATEEPTFIPEPWSACSTTCGPGVQVREVKCRVLLTFTQT 526

Qy        601 EIELPEEEECEGPKLPTERPCLLEACDESPASRELDIPLPEDSETTYDWEYAGFTPCTAC 660
              ||| ||||||||:| |||||| ||||||||||||||||||||||||||||.||||||:|| |
Db        527 EIELPEEEECEGPKLPTERPCLLEACDESPASRELDIPLPEDSETTYDWEYAGFTPCTAC 586

Qy        661 LGGHQEATAVCLHIQTQQTVNDSLCDMVHRPPAMSQACNTEPCPPR 706
              |||||||||:::||||||.||||||||||||||||||||||||||||
Db        587 LGGHQEATAVCLHIQTQQTVNDSLCDMVHRPPAMSQACNTEPCPPR 632


RESULT 15
AAB01431
ID   AAB01431 standard; Protein; 874 AA.
XX
AC   AAB01431;
XX
DT   20-OCT-2000  (first entry)
XX
DE   Human TANGO 224 (form 2).
XX
KW   TANGO; 128; 140; 197; 212; 213; 224; 239; modulating agent; asthma;
KW   graft versus host diseases; rheumatoid arthritis; psoriasis;
KW   inflammatory bowel disease; septic shock; ulcerative colitis;
KW   Crohn's disease; chronic myelogenous leukemia; cancer; liver
KW   disease; Hodgkin's disease; osteoarthritis; Lyme's disease;
KW   cachexia; autoimmune disease; myasthemia gravis; autoimmune diabetes;
KW   systemic lupus erythematosus; transgenic animal; diagnosis;
KW   prognosis; prophylatic; therapeutic; human.
XX
OS   Homo sapiens.
XX
PN   WO200039284-A1.
XX
PD   06-JUL-2000.
XX
PF   23-DEC-1999;   99WO-US30405.
XX
PR   30-DEC-1998;   98US-0222546.
XX
PA   (MILL-) MILLENNIUM PHARM INC.
XX
PI   Holtzman DA;
XX
DR   WPI; 2000-465743/40.
DR   N-PSDB; AAA47482.
XX
PT   Novel nucleic acid sequences encoding TANGO 128, 140, 197, 212, 213,
PT   224 and 239 polypeptides useful for the treatment of asthma, rheumatoid
PT   arthritis, psoriasis and autoimmune diseases
XX
PS   Claim 8; Fig 30; 209pp; English.
XX
CC   Nucleic acids encoding TANGO polypeptides are useful as modulating
CC   agents for regulating cellular processes like asthma, graft
CC   versus-host diseases, rheumatoid arthritis, psoriasis, inflammatory
CC   bowel disease, septic shock, ulcerative colitis, Crohn's disease,
CC   chronic myelogenous leukemia, cancer, liver disease, Hodgkin's
CC   disease, osteoarthritis, Lyme's disease, cachexia and autoimmune
CC   diseases e.g. myasthemia gravis, autoimmune diabetes and systemic
CC   lupus erythematosus. The nucleic acids are also useful for producing
CC   transgenic animals and the TANGO polypeptides themselves. Partial
CC   TANGO-128, 140, 197, 212, 213, 224, 239 sequences are useful in
CC   forensic biology, for diagnostic assays, prognostic assays,
CC   pharmacogenomics and for monitoring clinical trials. TANGO
CC   polypeptides are suitable for both prophylactic and therapeutic
CC   methods for treating a subject at risk of a disorder or having a
CC   disorder associated with aberrant TANGO expression. A wide range
CC   of cellular disorders can be treated.
XX
SQ   Sequence   874 AA;

  Query Match             30.9%;  Score 2891;  DB 21;  Length 874;
  Best Local Similarity   58.4%;  Pred. No. 3.5e-191;
  Matches  493;  Conservative  126;  Mismatches  195;  Indels   30;  Gaps    5;

Qy         64 TSRNTRSDEDKDGNWDAWGPWSECSRTCGGGASYSLRRCLSGKNCEGQNIRYKTCSNHDC 123
              :||   ||:||:|| ||||| ||:|||||||||||||||||||: ::|||:||||:||| ||
Db         22 SSRTARSEEDRDGLWDAWGPWSECSRTCGGGASYSLRRCLSSKSCEGRNIRYRTCSNVDC 81

Qy        124 PPDAEDFRAQQCSAYNDVQYQGHYYEWLPFRYNDPAAPCALKCHAQGQNLVVELAPKVLDG 183
              ||:| |||||||:|||::| :||||   |:|      |:|:| |:|   |||||||||||
Db         82 PPEAGDFRAQQCSAHNDVKHHGQFYEWLPVSNDPDNPCSLKCQAKGTTLVVELAPKVLDG 141

Qy        184 IRCNIDSLDMCISGICQAVGCDRQLGSNAKEDNCGVCAGDGSTCRLVRGQSKSHVSPEKR 243
              ||| |:|||||||||:|| |||| ||||   ||||||| |:|||:||||||| || :| |
Db        142 IRCYIESLDMCISGLCQIVGCDHQLGSTVKEDNCGVCNGDGSTCRLVRGQYKSQLSAIKS 201

Qy        244 EENVIAVPLGSRSVRIIVKGPAHLFIESKTLQGSKGEHSFNSPGVFVVENTIVEFQRGSE 303
              :: |:|:| ||| :|: :||| ||::|:||||||:|||:| :| | |:|:|:|:||:   :
Db        202 DDTVVAIPYGSRHIRLVLKGPDHLYLETKTLQGTKGENSLSSTGTFLVDNSSVDFQKFPD 261

Qy        304 RQTFKIPGPLMADFIFKTRYTAAKDSVVQFFFYQPISHQWRQTDFFPCIVTCGGGYQLNS 363
              ::  :: ||| :||| | | : : |: |. .|||| |:: :::||||: ||:||||: :
Db        262 KETLRMAGPLTADFIVKIRNSGSADSTVQFIFYQPIIHRWRETDFFPCSATCGGGYQLTS 321

Qy        364 AECVDIRLKRVVPDHYCHYYPENVKPKPKLKECSMDPC          PSSDGFKEIMPYD 414
              ||| |:|   ||| | ||||||||:||||||:||:::|||          |:|||:|:|||||
Db        322 AECYDLRSNRVVADQYCHYYPENIKPKPKLQECNLDPCPAQRVGTTVENSDGYKQIMPYD 381

Qy        415 HFQPLPRWEHNPWIACSVSSCGGIQRRSFVCVEESMHGEILQVEEWKCMYAPKPKVMQTC 474
               : ||||||   || ||| ||||||| |:   |:|| : | :    ||||||||| || :  | |
Db        382 LYHPLPRWEATPWTACSSSCGGGIQSRAVSCVEEDIQGHVTSVEEWKCMYTPKMPIAQPC 441

Qy        475 NLFDCPKWIAMEWSQCTVTCGRGLRYRVVLCINHRGEHVGGCNPQLKLHIKEECVIPIPC 534
              |:||||||:| ||| |||||:||||||||||:|||  | |||:|: | |||||||::| ||
Db        442 NIFDCPKWLAQEWSPCTVTCGQGLRYRVVLCIDHRGMHTGGCSPKTKPHIKEECIVPIPC 501

Qy        535 YKPKEKSPVEAKLPWLKQAQELEETRIATEEPTFIPEPWSACSTTCGPGVQVREVKCRVL 594
              |||||| ||:||||| |||||||        :|||:|||| |||:|||: ||| :||:||
Db        502 YKPKEKLPVEAKLPWFKQAQELEEGAAVSEEPSFIPEAWSACTVTCGVGTQVRIVRCQVL 561

Qy        595 LTFTQTEIELPEEEECEGPKLPTERPCLLEACDESPASRELDIPLP          EDSET 644
              |:|:|:  :||| :||||||   ::|      ||  : : :      |          :| :
Db        562 LSFSQSVADLPIDECEGPKPASQR---ACYAGPCSGEIPEENPDEIDGLEGGLQDFDE 616
```

```
Qy        645 TYDWEYACFIPCIAICLCGHQEAIAVCLHIQTQQIVNDSLCDMVHRPPAMSQACNTEPCP 704
               ||||| |||  |:  :|  ||  |||:   ||: ||::     ::||      |||  :  ::||  :|||
Db        617 LYDWEYEGFIKCSESCCCGVQFAVVSCLNKQTREPAEENLCVISRRPPQLLKSCNLDPCP 676

Qy        705 PRWHVGSWGPCSATCGVGIQTRDVYCLH-----PGETPAPPEE-CRDEKPHALQACNQFD 758
               ||  :|  |  |||  ||  ||:|,|||:| |          ||      :|  ||   ||   :||||:|:
Db        677 ARWEIGKWSPCSLICGVGLQTRDVFCSHLLSREMNETVIIADELCRQPKPSTVQACNRFN 736

Qy        759 CPPGWHIEEWQQCSRTCGGGTQNRRVTCRQLLTDGSFLNLSDELCQGPKASSHKSCARTD 818
              |||  |:    :||  ||||||||  |  |  |  |:|  :  |||||  |  :    |     |  :    ::|  :  |
Db        737 CPPAWYPAQWQPCSRTCGGGVQKRRVLCKQRMADGSFLELPETFCSACKPACQQACKKDD 796

Qy        819 CPPHLAVGDWSKCSVSCGVGIQRRKQVCQRLAAKGRRIPLSEMMCRDLPGFPLVRSCQMP 878
              ||       :  ||::||  .||  |  |  |  |    :|:::     |        ::    :|   ||       :|  |  :
Db        797 CPSEWLLSDWTECSTSCGEGTQTRSAICRKMLKTGISTVVNSTLCPPLPFSSSIRPCMLA 856

Qy        879 ECSK 882
                |::
Db        857 TCAR 860
```

Search completed: March 22, 2003, 13:28:20
Job time : 72 secs

```
                         GenCore version 5.1.4_p5_4578
                  Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on:         March 22, 2003, 13:22:39 ; Search time 39 Seconds
                                           (without alignments)
                                           4168.290 Million cell updates/sec

Title:          US-09-784-358-2
Perfect score:  9346
Sequence:       1 MASWISPWWVLFGMVFMHSP..........LNLSLDKYKQRSQSCQISI 1691

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters:       283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :        PIR_73:*
                1:   pir1:*
                2:   pir2:*
                3:   pir3:*
                4:   pir4:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                SUMMARIES
                    %
Result              Query
  No.    Score      Match Length DB  ID                             Description
                                                                     -
   1    1183.5   12.7   2165  2  T21371                       hypothetical prote
   2    1110.5   11.9   1059  2  T22545                       hypothetical prote
   3       863    9.2   1558  2  C89114                       protein C37C3.6a [
   4       863    9.2   2167  2  T34395                       hypothetical prote
   5       700    7.5    951  2  T00260                       hypothetical prote
   6     593.5    6.4   1205  2  T18517                       procollagen N-endo
   7       593    6.3    808  2  T14764                       hypothetical prote
   8     571.5    6.1    550  2  T47158                       hypothetical prote
   9     564.5    6.0    951  2  T00017                       gene ADAMTS-1 prot
  10       490    5.2   1444  2  T18856                       angiogenesis inhib
  11       415    4.4    837  2  T00355                       hypothetical prote
  12       352    3.8    957  2  T15976                       hypothetical prote
  13     341.5    3.7    788  2  T25061                       hypothetical prote
  14     292.5    3.1   1522  2  T00028                       brain specific ang
  15     289.5    3.1   1584  2  T00026                       brain specific ang
  16       276    3.0    803  2  A47723                       F-spondin precurso
  17       274    2.9    437  2  S05478                       properdin - mouse
  18       271    2.9    469  1  S29126                       properdin precurso
  19       267    2.9    904  2  T00326                       hypothetical prote
  20       265    2.8    807  2  A38152                       F-spondin - rat
  21       261    2.8   5175  2  T20992                       hypothetical prote
  22     257.5    2.8   1074  2  JC5928                       semaphorin F precu
  23     256.5    2.7   3707  2  S18252                       heparan sulfate pr
  24     254.5    2.7   5198  2  T43290                       hemicentin precurs
  25     253.5    2.7   1360  2  T33922                       hypothetical prote
  26       253    2.7   1896  2  T08851                       Down syndrome cell
  27     249.5    2.7   1572  2  T00027                       brain specific ang
  28       247    2.6    712  2  A45638                       immunodominant mic
  29     242.5    2.6   1178  1  A39804                       thrombospondin pre
  30     241.5    2.6   4162  2  T42633                       connectin/titin
  31       241    2.6   5376  2  T42215                       zonadhesin - mouse
  32       240    2.6   1170  1  TSHUP1                       thrombospondin 1 p
  33       240    2.6   1172  1  TSHUP2                       thrombospondin 2 p
  34     237.5    2.5   4391  2  A38096                       perlecan precursor
  35     237.5    2.5   7962  2  T38346                       elastic titin - hu
  36     236.5    2.5    724  2  A48569                       antigen Em100 - Ec
  37       235    2.5   1172  2  A42587                       thrombospondin 2 p
  38       234    2.5   1170  2  A40558                       thrombospondin 1 p
  39       233    2.5    805  2  T34212                       hypothetical prote
  40     230.5    2.5   1344  2  T14316                       rig-1 protein - mo
  41     229.5    2.5   6642  2  T29757                       protein UNC-89 - C
  42     220.5    2.4   1188  2  D86236                       protein F14N23.5 [
  43       215    2.3    860  2  T16892                       hypothetical prote
  44     213.5    2.3   1033  2  S19247                       cell adhesion prot
  45     213.5    2.3   1272  2  S26180                       neurofascin - chic

                                ALIGNMENTS

RESULT 1
T21371
hypothetical protein F25H8.3 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999
C;Accession: T21371; T24896
R;Gajadsty, S.
submitted to the EMBL Data Library, February 1996
A;Reference number: Z19413
A;Accession: T21371
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-2165 <WIL>
A;Cross-references: EMBL:Z69360; PIDN:CAA93287.1; GSPDB:GN00022; CESP:F25H8.3
A;Experimental source: clone F25H8
R;Gajadsty, S.
submitted to the EMBL Data Library, February 1996
A;Reference number: Z19949
A;Accession: T24896
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-2165 <W12>
A;Cross-references: EMBL:Z69361; PIDN:CAA93288.1; GSPDB:GN00022; CESP:F25H8.3
A;Experimental source: clone T13H10
C;Genetics:
A;Gene: CESP:F25H8.3
A;Map position: 4
A;Introns: 31/1; [illegible]

  Query Match             12.7%;  Score 1183.5;  DB 2;  Length 2165;
  Best Local Similarity   22.7%;  Pred. No. 5.2e-62;
  Matches  402;  Conservative  212;  Mismatches  576;  Indels  580;  Gaps   80;

Qy         75 DGNWDAWGDWSDCSRTCGGGASYSLRRCLI        GRNCGGQNIRYKTCSNHDCPPDAE 128
              || |   |   | :||||||||      || |  :           |: | ||   ||::|:  :|| | :
Db        603 DGQWGDWRSWGECSRTCGGGVQKGLRKCDSPKPRNGGKYCVGQRERYESCNTQECTWDIQ 662

Qy        129 DFRAQQCSAYN  DVQYQG        HYYEWLPRYND  PAAPCALKCHAQGQNLVVELAP 178
               :|  ||  :|   : :  |:         |     |:|:| :  |    | | |     |        |
Db        663 PYREVQCSEFNNKPFTGQCVAGINH    WVPKYANVAINEPCKLYCHSCNAFYLLPG 719

Qy        179 KVLDGTRCNTDSLDMCISGICQAVGCDRQLGSNAKEDNCGVCAGDGSTCRLVRGQSKSHV 238
              ||:|||  |:  :   |:|::| |    ||| || |   :  |  ||||  ||  |:|::|:|
Db        720 KVVDGTPCDRNGDDICVAGACMPAGCDHQLHSTLRRDKCGVCGGDDSSCKVVKGTFNEQG 779

Qy        239 SPEKREENVIAVPLGSRSVRITVKGPAHLFIESK  TLQGSKGEHSFNSPGVFVV      291
               :     |    |: :| || ::  |   ||   : :   :        :|: : ||      :  | | |
Db        780 TPGYNL  VMKIPAGSANIDIRQKGYNNMKEDDNYLSLRAANGEFLLN   GHFQVSLARQ 835

Qy        292 -    ENTIVEFQRGSERQIEKI  PGPLMADF-  -          ITKIRYIAAKDSVVQ 332
                   ::| :|:   ||:      :|     ||: :|                    |      |||   : |
Db        836 QIAFQDIVLEYSGSDATIERINGIGPIRSDIYVHVLSVGSHPPDISYEYMIAAVPNAV  895
```

```
Qy        333 FFFYQPIS---HQWRQTD-FFPCTVTCGGGYQLNSAECVDIRLKRVVPDHYCHYYPENV- 387
                  :|||     : || || :   |    |   | |      |:|:    |    |   |    :||
Db        894 ---IRPISSALYLWRVIDTWTECDRAC-RGQQSQKLMCLDMSTHRQSHDRNC----QNVL 945

Qy        388 EKEKEKEKESMD---------           --PCPSSDGFKEIMPYDHF----- 416
              |||    : |::|                            |    :| ::      :|
Db        946 KEKQATKMCNIDCSIKWIIEDVSSCSAKCGSGQKKQRVSCVFMEGDKQTFASEHLCDKNS 1005

Qy        417 QPLP    .      RWEHNPWTACSVSCG--GGIQRRSFVCVEESMHGEIIQVEEWKCM 463
              :|              :|  :    ||:||  :||    |  :  |:|:  ||::|         :|:|    |
Db       1006 KPSDIASCYIDCSGRKWNYGEWTSCSETCGSNGKMHRKSY-CVDDSNR----RVDESLCG 1060

Qy        464 YAEEKVMLTINLEDCPEWIAMEWSQTVIGCRGLRYRVVICIN -HRGEHVGGCNP-QL 520
                   |      :  ||      ||:|:      ||:|:  :|     |::  |      |::     |    |     | | |
Db       1061 REQKEATERECNRIPCPRWVYGHWSECSRSCLGGVKMPHAQCLDAADRETHTSRCCPAQT 1120

Qy        521 ELHIKEFTVIPIPCYKIEEESIVEAKLIWLKQAQELEETRIAIEEIIFIPEEWSACSITC 580
              : |    |                           | :                               || ||   |
Db       1121 QEHCNEHA-------------------CTWWQFG------------------VWSDCSAKC 1144

Qy        581 GPGVQVREVKCRVLLTPTQTETELPEEECEGPKLPIERPCLLE----------------AC 625
              | ||| |:   |              :   : |||   |     :      :||   |                   :|
Db       1145 GDGVQYRDANCI----DRHRSVLPEHRCLKMEKIIIKPCHRESCPKYKLGEWSQCSVSC 1199

Qy        626 DESPASR-----------ELDIPL-------PEDSETTY-----DWEYAGFIPCIATCLGG 663
              ::   :||            |:|:  |           |      :|              |      :: |  :|   |
Db       1200 EDCWSSRRVSCVSGNGTEVDMSLCGIASDRPASHQICNLGICPFWRNTDWSACSVSCGIG 1259

Qy        664 HQEAIAVCLHIQTQQTVNDSLCDMVHRPPAMSQACNTEPCPPRWHVGSWGPCSATCGVGI 723
              |:|      |::      :|:|:  |  |          :  |   || |:   ||        |           | |||  ||| ||
Db       1260 HRERTTECIY--REQSVDASFCGDT-KMPETSQICHLLPC-TSWKPSHWSPCSVTCGSGI 1315

Qy        724 QTRDVYCLHPGEIPAPPEE-C-RDEKPHALQACN-------------QFDCPP-CWHIEE 767
              ||| |  |      |         |   |  | |:  :|      :  |                    | |  ||    |
Db       1316 QTRSVSCTRGSEGTIVDEYFCDRNTRPRLKKICEKDTCDGPRVLQKLQADVPPIRWATGP 1375

Qy        768 WQQCSRICGGCTQNRRVICRQLLTDGSPLNISDFLCQG PKASSHKFCAKISCPHHLAVG 826
              |  || ||| ||| |  :  ||    : |           | || |         |    |  ::|        ||    :    :
Db       1376 WTACSATCGNGIQRRLLKCRDHVRD-----LPDEYCNHLDKEVSTKNCRLEDC-SYWKMA 1429

Qy        827 DWSKCSVSCGVGIQRRKQVCQRLAAKGRRIPLSEMMCRDLPGFPL VRSCQMPLCSKIKS 885
              :|  :|   :||   :|:  :  |       |    |    |   |  ::  |  |:       |      |:|::   |  |
Db       1430 EWEECPATCGTHVQQSRNVICVSAEDGGRFILKDVDC-DVQKRPISARNCRLEPCPK-   1485

Qy        886 EMKIKLGEQCPQILSVQPVYIQTRPTKRINLTIGCRAYLLFNTSVIIKCPVKRFQKSLIQ 945
                        ||:                                    |||   ::: :  |      ||
Db       1486 ------GEE------------------------HIGS--WIIGDWS---KC----------- 1501

Qy        946 WEKDGRCLQNSKPLCITKSGSLKIHGLAAFLIGVYRCIAGSAQLIVVLKLIGTDNRLIAR 1005
                       |       :|    ::                                    |  :  |    ||     |:      |    :
Db       1502 ---SASCGGGWRRRSVS-------------------CTSSSCDETRKPKMFDKCNEELCP 1539

Qy       1006 PALPEPMPEYPGMDHSEANSLGVTWHKMRQMWNNKNDLYLDDDHISNQPFLRALLGHCSN 1065
              |      :   |                               |                               |||
Db       1540 PLTNNSWQISP-----------------------WT------------------------HCSV 1556

Qy       1066 SAGSTNSWELKNKQFEAAVKQCAYSMDTAQFDELIRNMSQLMEIGEVSDDLASQLIYQIV 1125
              |  |                    |                     ::::         :|     ||:
Db       1557 SCG-------------------GGVQRRKIWCEDVL---------SGRKQDDI--------- 1581

Qy       1126 AELAKAQPTHMQWRGIQEETPPA-AQLRGETGSVSQSSHAKNSGKITEKPKG---PVLMR 1181
                |  ::  :|     :         | ||   :        :| |  |  :|  :  ::      |           |:|
Db       1582 -FCSFIKPREQR----DCRMPPCRSHYHNKTSSASMISLSSSNSNTTSSASASSLPIL-- 1634

Qy       1182 QSQPPTTSPN      KTINPKIGNIVEIIKR   IEVINILCDLIIPSEAIYIWIKDGILLQ 1235
                    || :|:            ::: |          |||    | :|
Db       1635 ---PPVVSWQTSAWSACSAKCGRG---TKRRVVECVN  ..    .      ----------- 1665

Qy       1236 PSVKIILDGTGKIQIQNPTRKEQGIYECSVANHLGSDVESSSVLYAEAPVILSVERNITK 1295
              ||:  :  :     |      |  :  |    :
Db       1666 PSLNVTVASTECDQIKKPVEE----        -----    .     ----------------- 1686

Qy       1296 PEHNHISVVVGGIVPAALGANVIIECPVKCVPQ-PNIIWIKRGGSIPSGNVSLLPNGSLLL 1354
                                        :||     |     |:          ||
Db       1687 -----------------  ---------VRCRIKHCPRWKITIWSS----------------  ----- 1704

Qy       1355 QNVSLPNPGTYVCIATNAIGKAVATSVIHIIEKRWPESPKIVFLQCHKKYILQA-TNTRIN 1413
                            |    |    |                 :  ||       :        :| |   |  :   :    |   :|
Db       1705 -------------CSVTCGRG----   -----IRRR----FVQCYRCRKNLVSDSPCNPKTK 1738

Qy       1414 SNDPTGEPPPQEPF--WEPGNWSHCSATCGHLCARIQRPQCVMANGQPVSEALCDHLQKP 1471
               |        |      |     |        |  |   |          :  :|     |:      :|:     :       :|:     :   |
Db       1739 LNSVANCPPVACPAYRWNVTPWSKCKDECAPGQKQTRRVHCISTSCKRAAPRMCELARAP 1798

Qy       1472 LAGPEPCNIEDCPAPWFTSVWSQCSVSCGPGYHSRQVTCKR IKANCIVQVV    SPRAC 1527
                            |:     :||    |          |     ||  |||||     :|:|    |:|            |   |:  ::
Db       1799 -TSIRECDTSNCPYEWVPGDWQICSKSCGPGVQTREVRCRRKINPNSTIPIIPMLEDEPA 1857

Qy       1528 APKDRPLGRKPCPGHPCVQWPPGNRCPGRCMCRAVPMQQRHTACQHNSSPSNCPDRKKPT 1587
               ||::                                                                  |:       :|
Db       1858 VPKEK------------------------------------  ------------CELFPKPN 1870

Qy       1588 LRRNCTSGACD--VCWHTGPWKPCTAACGRGPQSRKVPCIRTKSCPPVAPRHCVQKKKPI 1645
                 :  |        ||        |   |||    |:     ||:|  :  |:|  |:    :  |   :   |      |||
Db       1871 ESQTCPLNPCDSEPKWSPCPWGECSKNCGQGIRRRSVKCV ANDGRRVPRVKCT IKKIK 1928

Qy       1646 SWRHCLGPSCDRDCTDTTHYCMFVKHLNLCSLD 1678
                 ::|         :|:|   :|    |    :|    |:  :  |
Db       1929 KTQYCP    -ERNCLPST--CQPLKSQNVKAKD 1955


RESULT 2
T22545
hypothetical protein F53B6.2 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T22545
R;White, S.
submitted to the EMBL Data Library, October 1996
A;Reference number: Z19578
A;Accession: T22545
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1059 <WIL>
A;Cross-references: EMBL:Z81086; PIDN:CAB03121.1; GSPDB:GN00019; CESP:F53B6.2
A;Experimental source: clone F53B6
C;Genetics:
A;Gene: CESP:F53B6.2
A;Map position: 1
A;Introns: 38/3; 62/3; [illegible]

  Query Match             11.9%;  Score 1110.5;  DB 2;  Length 1059;
  Best Local Similarity   21.1%;  Pred. No. 4.4e-58;
  Matches  351;  Conservative  170;  Mismatches  451;  Indels  695;  Gaps   52;

Qy         77 NWDAWGDWSDCSRTCGGGASYSLRRCLTGRNCEGQNIRYKTCSNHDCPPDAEDFRAQQCS 136
              :| ||   || |::||||| |   ||||| : | |:::|:| |:    |    :    |    |
Db         25 SWAAWSPWSSCTKICGGGVSRQLRRCLTSK-CSCESVRPKVCAQKTCPSKSRLARDICG 83

Qy        137 AYNDVQYQGHYYEWIPPYNDPAAPCALKCHAQ--GQNLVVELAPKVLDGIRCN IPSLDM 193
                  ::   :|                 | : | ::  | |  :          :| |||| |       |   :
Db         84 G-EEIVSRGQ---------------CEVVCRSRLTGANFLW-  --RVDDGTPCQAATSRAV 124

Qy        194 CISCICQAVGCDEQLGSNAKEENCGVTACLCAICRLVRGQSASHVSPLKKREENVIAVFLG 253
               |  | ||  ||||   :  |:  :  |  ||||  |  |  ||
Db        125 CSKGSCQIVGCDGLISSSFRFDACGVCGGRGDTC------- -----     ------  -- 158

Qy        254 SRSVRITVECPAHLPIDSPILQLKELHLINSGVEVVEKIIVLIQRGSERQIEKIPGPL 313
                                                            :|
Db        159 -----------------  -----  ---------------DN------  -----  --  -- 160

Qy        314 MADFIFKT   RYTAAK---DSVVQPPYQPISHQWRQTDFFPCTVTCGGGYQLNSAE--- 365
                   ||:|      |||     | :|               |                     | |  :  |
Db        161 -GKFIWKVSEEYTACASNCDDIV-------  --DW   ------  --SGAGRSIASTSQPI 198
```

```
Qy        364 CVLCELKRVVPCHYCHYYPENVKPKPKLKECMHPCPSCPCKKCIMPYCHPQPLPEWE 423
                 ||:     ||||    |         : ::||  :  :  |  |   |||              |:
Db        199 VVCVNAIIGRVVPEKIC    ADKLPPKVEAPPCPMLICPS--------------PWM 238

Qy        424 HNPWIACSVSCGGGIQRRSPVCVEESMHGEILQVEEWKCMYAPKPKVMQICNLPDCPKWI 483
                 |:   |    || |  ::|     :|  : :  |      :  |  :  |       ::    :  |     |  :|
Db        239 AADWIECVPHCGEGIRKKREVYCV QIAHNVIVHVPDTPCENGIRPAAEENCVSTSCGRWE 297

Qy        484 AMEWCQCIVICCECCLRYPVVLCNHRGEHVCG   CNPQLKIHDKED V       CKEKLKS 541
              |  :||  ||  :||:|:|  |  |  |  |          |||   |:    :    :   |     :||        |
Db        298 AGKWSKCIASCGQGVRPRHVAC           VGGSDCDEGGRPRQEIICYAGIPC        S 344

Qy        542 PVPAKIPWLKQALELEEIPIAI          EEPIPIPEPWSACSIICGPGVQVREVKCRVL 594
                        |  |   :|    |:       .          :  |   :    ||  ||:|||  ||    |  |:|    :
Db        345 IAINSIDWNDRAY LSGNIEGSMPNHNDWQAPRLVAGEWSPCSSTCGHGVMSRTVECVAV 403

Qy        595 LTPIQIEIELPPEECEGPKLPIERPCLLEACDESPASRELDIPLPEDSEITYDWEYAGFI 654
                      :        :||   |::        |:|  |  |:|:           :  |    |  |  |    :     |   :|
Db        404 NPISSAPIKLPMSECQ      DQEQPKLPESCEVESCPLQEDSKLSED EAPYQWKYGDWI 458

Qy        655 PCIAICGGHQEAIAVCLHIQTQQIVNDSLCDMVHRPPAMSQACNIEPCPPRWHVGSWGP 714
                 |:|:||||  |:|      |:  :  :  ::|    |  ||       |||    |:  ||    ||||  |         :
Db        459 QCSASCGGGKQKAALKCIQVSIGKSVQWSQCPAPKKPPEKSRPCNQHPCPPWLTSKYGC 518

Qy        715 CSAICGVGIQTRDVYCLH               PGEIPAPPEECRDEKPHALQACNQFDCPPGWHIE 766
              ||  :||  |    |  |  |                             :  |    :||      :  ||     ||   |
Db        519 CSMSCGSGIAPRSVKCAQIVSKIIGADAHIVLKDSRCHRKKPQIEII NVVACPAIW     575

Qy        767 EWQQCSRTCGGGTQNRRVICRQLLTDGSFLNLSDELCQGPKASSHKSCARTDCPPHLAVG 826
                                 : |:|                                 |     :|               |
Db        576 -          VSCCNKRRN                            RIRINR               CPIA 591

Qy        827 DWSKCSVSCGVGIQRRKQVCQRLAAKGRRIPLSEMMCRDLPGFPLVRSCQMPECSKIKSE 886
                 |::||  ||   |  :|!:   ::      ::|:         ::|                 |    |         :|:
Db        592 QWIECSRSCDSGERRRQVWCEIRLSPGKIQRRPDVECDANIKPQIVEVCSGNSCSR      650

Qy        887 MKIKLCPLCPLTTCVLPVYILTPTDKPINLFICTRAYLLPNICVIIKCCVCIIQKCCIW 946
                                 |::||    ||:  :    :|:|:  |  ||    |  |      ||:  ||||  ::|  |     |
Db        651                    PELLS NPVPTLNAPLKKLIICTCVAIIYLCIICIKCKCAKRICKKKCIW 700

Qy        947 EKDGPCIQNSKRLGIIKSGSLKIHGLAAPIIGVYRCIAGSAQEIVVLKLIGIDNRLIAKP 1006
               :|:|:  :||       :  :  :  :  .  ::              ||||  :         |   |  |
Db        701 KKNGKKIKNPAIIKVCANCNIPVLHARMLDAGVYICTIGRLQSNVILNR            749

Qy       1007 ALRPPMKRYPCMSHCEANCLCVIWHKMPQMVNNKNLLYLIGCHIICNQPTLLALLGHCSNS 1066
                        |::!
Db        750       KYRGFP                                                                 755

Qy       1067 ACSINSWELKNKQFPAAVKQIAYCMIIALPTGLIPNMCQMLICHVGGLAGQLTYQIVA 1126

Db        756                                                                               755

Qy       1127 ELAKAQPTHMQWRCIQPPTPPAAQLRGKIGGVSQGSHAKNSGKLIIKPKCPVLMRQSQPP 1186
                                                                     : ::     |:             :
Db        756                                   -   -ASRVDLAPK-     --    -- P 765

Qy       1187 SISPNKIINGRIGNIVYIIKRIEVINILCDLIIPSEAIYIWIKDGTLLQPSVKIILDGTG 1246
              |                                         ||        |::      :|   |  :  :|
Db        766 QI                                        PS        IKNRQPVQVSKEDVLR     785

Qy       1247 KIQIQNPIRKEQGIYECSVANHLGSDVESSSVLYAEAPVIISVEPNIIKPEHNHLSVVVG 1306
                                                   |  :|||:                 :  |::
Db        786                                       EQASVLH               KMNVS               797

Qy       1307 GIVEAALGANVTIRCPVKGVPQPNITWLKRGGSLSGNVSLLPNGSLLLQNVSLENEGTYV 1366
               ::||  |  |                    .         :          ||
Db        798  LIEALLTA                 PNDEKAREQLRKYGN      -                    -   820

Qy       1367 CIAINALGKAVAISVLHLLEPPWPESRIVPLQGHKKYIIQAINTRINGNDPIGEPPPQEP 1426
                                 |:  :
Db        821                  ELVAR                                                            825

Qy       1427 PWEPGNWSHCSATCGHGARTQRPQC VMANGQ   EVSEALCDHLQ   KPLAGFEPCNI 1480
```

```
               |: |:||  |     :|     |  :  :     |  |    :::      |    :.:.   |       :|         ||:
Db        826 WDIGHWSECRQKTCHVAGYQARGISCKVIFHGEIRNVDNSICESIASVRPCPEIRPCHR 883

Qy       1481 RDCPARWFISVWSQCSVSCGEGYHSRQVICKRIFANGIVQVVSPRACAPKDRPLGRKPCF 1540
               |||  ||    |  ||:||
Db        884 EDCP-RWEASQWSECSSQ----------------------------------------   -    900

Qy       1541 GHPCVQWEPGNRCPGRCMCRAVEMQQRHIACQHNSSDS         NCIGRKKRPIIRRNCTSGA 1596
                                   ||:     :    ::|:   |:    :      |      :||     ||         :|  :
Db        901         -       -        RCVSSMLAQKRENVICKPINGIGVLIQHCIGIINRPAIIMICPNQN 945

Qy       1597 CDVCWHTGPWKPCTAACGR GPQSRKVDCIHIRSCKPVAKRHCVQKKKPISWRHCLG       1652
                   |      |   |    |     |::  ||    |  |  |  :  |:       |  :|       |:|  |  ::|  ,  |  |:
Db        946 CKAEWRTSDWGSCSSECGIGVQLRELSCVWISSGRP AGRNCEQMRRPHSARACVADEP 1004

Qy       1653  PSC            -        DRDCIDTIHYCMPVKHLNDCSLDRYKQRCCQSC 1688
                 |  |                   |    |   |   :  :|    :|    :  |          :|:||  :|
Db       1005 LPPCMPIASALYQEDASCQDQSRPCDITKLPHSCGSLEVRQKCCSIC 1051


RESULT 3
C89114
protein C7C3.6a [imported] - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 18 May 2001 #sequence_revision 18 May 2001 #text_change 18 May 2001
C;Accession: C89114
R;anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A;Title: Genome sequence of the nematode C. elegans: a platform for investigating bio
A;Reference number: A75000; MUID:99069613; PMID:9851916
A;Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_
A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999;
A;Accession: C89114
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1558 <STO>
A;Cross-references: GB:[illegible] GSPDB:GN00023; [illegible]
C;Genetics:
A;Gene: [illegible].6a
A;Map position: 5

  Query Match                9.2%;  Score 863;  DB 2;  Length 1558;
  Best Local Similarity    28.6%;  Pred. No. 4.1e-43;
  Matches  220;  Conservative   14;  Mismatches  412;  Indels  122;  Gaps   30;

Qy         35 LPEEFALSPQGSFLEDIIGEQFL                          IYRYDQISENIRSDED 73
              :| :|  |       |    |    . :|                           |    |  :    :|  :
Db         15 VPIWAFSLSSFFGSDVAQKPYIHPNSPPERDPASSPMKRQAYQVYVPGDVSVIVDKSDQK 74

Qy         74 KDGNWDAWGDWSDCSRTCGGGASYSLRRCLIGRNCEGQNIRYKICSNHDCTPLAEDFRAQ 133
              :  |||    |      ::|||:|||'          :|       :|  |  :::||  :|: :  :     ||||:
Db         75 EIGNWGPWVPENDCSRSCGGGVQLEKRQC   SGLCIGASVRYIGCNLNACIESGIDFRAE 131

Qy        134 QCSAYNDVQYQGHYYEWLPRYNDPAAPCALKCHAQGQNLVVELAPKVLDGTRCNTDSLDM 193
              ||| :||       |:|::| |  |             |  |  |    :    |         :  '  ||:|||:|::  |  |:
Db        132 QCSKENDEALDGNYHKWIP--YKGKNKCELVCKPESGNPYYKWAPKVVPGIKCGSKSNDI 189

Qy        194 CISGICQAVGCDRQLGSNAKEDNCGVCAGDGSTCRLVRGQ SKSHVSPEKREENVIAVPL 252
              |:  |  |    ||||  :|||:  |  |  ||  |  ||||||:  :  |:    :  ::||         :    :|  :|
Db        190 CVDGECLPVGCDGKIGSSLKEDKCGKCGDCGSICKEIIGRPDERNLSPGYHI      IIKIPE 247

Qy        253 GSRSVRITVKGPAHLFIESKTLQGSKGEHSFNSPGVFVVENT          IVEFQRGSERQTF 307
              :: :::|           '           :    ,:                  |    |:     |                  :         :|   ::
Db        248 GAINIKIQ      EARKSINNLALKNGSGHPYLNGNGLIQVEKEVEVGGIIEVYDDAEPEIL 304

Qy        308 KIPGPLMADFIIKIRY TAAKDSVVQFFFYQP         IGHQWRQTDPTPCTVTCGGGYQLN 362
                    |||      :                 :          ::|:  :::  |     |                :  :   ::   ::   ||:|:||   |   |
Db        305 SAQGPLSEELIVALLFRKGSRGIAIKYEFSIPIELEVIYMYKEINWIPCSVSCGKGVQIR 364

Qy        363 CARCVDIRLKRVVPCHYCHYYPENVKPKPKLKECSMEPCPSCDGKKCIMPYCHPQPLPK 421
              :      |:|    :   |      |    |      ||          ||:     |  |          :  :
Db        365 NLYCIDGKNKGRVEDDQCE       PNNAIKPEEEKSCEIVDCLA                          E 404
```

```
Qy        422 WEHNPWTACSVSCGG-GIQRRSFVCVEESMHGEILQVEEWKCMYAPKPKVMQTCNLFDCP 480
              |      | :|| :||  | | |     | :      :|  : ||:   |       :| |  |||| | ||
Db        405 WFTGDWESCSSTCGDQGQQYRVVYCHQVFANGRRVTVEDGNCT-VERPPVKQTCNRFACP 463

Qy        481 KWIAMEWSQCTVTCGRGLRYRVVLCINHRGEHVGGCNPQLKLHIKEECVIPIPCYKPKEK 540
              :| |   || |:   ||     :|| | |      | | |       : ||    : |
Db        464 EWQAGPWSACSEKCGDAFQYRSVTC---RSEKEG---EEGKLLAADAC------------ 505

Qy        541 SPVEAKLPWLKQAQELEETRIATEEP----TFIPEPWSACSTTCGPGVQVREVKCRVLLT 596
               | :          : :: :  |        |     ||:     |: | | |       : ||| |:
Db        506 -PAD-------EQEKFDTERTCNLGPCEGLTFVTGEWNLC-TRCNDTEETREVTCK---- 552

Qy        597 FTQTETELPEEECEGPKLPTERPCLLEACDESPASRELDIPLPEDSETTYDWEYAGFTPC 656
                           : :|    | |: ||::    | |              |         |:|  : :: |
Db        553 ------------DSQGRAYPLEK-CLVDNSTEIPTDTPSCATQP---PCEYEWTVSEWSKC 597

Qy        657 TATCLGGHQEAIAVCLHIQT--QQTVNDSLCDMVHRPPAMSQACNTEPCPPRWHVGSWGP 714
              |   |   ||:        :|      |        : |::    |        :|      :      | | |      |:   ||
Db        598 TTECGHGHKTRRVICAIHQNGGLEVVDEGHC-QAEKPEGKINCINEEKCIGIWYTSSWSE 656

Qy        715 CSATCGVGIQTRDVYCLHPGETPAPPEECRDE--KPHALQACNQFDCP 760
              |:| || | | |     ||:    : |   || | || ||    | ||   |||
Db        657 CTAECGGGSQDRVAVCINYDKKPV-PEWC-DEAVKPSEKQDCNVDDCP 702
```

RESULT 4
I34395
hypothetical protein C37C3.6b - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29 Oct 1999 #sequence_revision 29 Oct 1999 #text_change 01-Dec-2000
C;Accession: T34395; T34394
R;Geisel, C.; Bradshaw, H.
submitted to the EMBL Data Library, July 1996
A;Description: The sequence of C. elegans cosmid C37C3.
A;Reference number: Z21518
A;Accession: T34395
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-2167 <GEI>
A;Cross references: EMBL:U64857; PIDN:AAC25868.1; GSPDB:GN00023; CESP:C37C3.6b
A;Experimental source: strain Bristol N2; clone C37C3
A;Accession: T34394
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1555,'SKF' <GE2>
A;Cross-references: EMBL:U64857; PIDN:AAC25867.1; GSPDB:GN00023; CESP:C37C3.6a
A;Experimental source: strain Bristol N2; clone C37C3
C;Genetics:
A;Gene: CESP:C37C3.6b; CESP:C37C3.6a
A;Map position: 5
A;Introns: 32/3; 104/2; 156/1; 207/1; 459/2; [illegible]; [illegible]; 1105/3; 1367/1; 1438/1; 1556/

```
  Query Match             9.2%;  Score 863;  DB 2;  Length 2167;
  Best Local Similarity   28.6%;  Pred. No. 6.4e-43;
  Matches  226;  Conservative  114;  Mismatches  312;  Indels  122;  Gaps   30;

Qy         35 LPEFALSPQGSFLEDTTGEQFL---------------------TYRYDDQTSRNTRSDED 73
              :| :| |      |  |    : :|                         |   | :     :| :
Db         15 VPTWAFSLSSFFGSDVAQKPYLHPNSPPERDPASSPMKRQAYQVYVDGDVSVTVDKSGQK 74

Qy         74 KDGNWDAWGDWSDCSRTCGGGASYSLRRCLTGRNCEGQNIRYKTCSNHDCPPDAEDFRAQ 133
              : |||   |     ::|||:||||        |:|      :| | ::|| :|: : |        ||||:
Db         75 ETGNWGPWVPENECSRSCGGGVQLEKRQC--SGDCTGASVRYISCNLNAC-ESGTDFRAE 131

Qy        134 QCSAYNDVQYQGHYYEWLPRYNDPAAPCALKCHAQGQNLVVELAPKVLDGTRCNTDSLDM 193
              ||| :||     :::::|  |          | | |    :    |      : | ||:|||:|:: | |
Db        132 QCSKFNEEALECNYHEWIP-YKC-ENKCELVCKPESCNFYYKWADKVVDGIKCDSKSNDI 189

Qy        194 CISGICQAVGCDRQLGSNAKEDNCGVCAGDGSTCRLVRGQ-SKSHVSPEKREENVIAVPL 252
              |: |  |   |||| :|||: | | || | ||||||: : |:    : ::||       :   :| :|
Db        190 CVDGECLPVGCDGKLGSSLKFDKCGKCDGDGSTCKTIEGRFDERNLSPGYHD--IIKLPE 247

Qy        253 GSRSVRITVKGPAHLFIESKTIQGSKGEHSFNSPGVFVVENT-----TVEFQRGSERQTF 307
              |: :::|       |        :  |:                       |  |:  ||       |:        :| :|
Db        249 CATNIKIQ----FARKCTNNIALENCSDHHYLNGNGLIQVEKEVEVGGTIEVYDDAEPETL 304

Qy        308 KIPGPLMADFIFKTRY-TAAKDSVVQFFFYQP----ISHQWRQTDFFPCTVTCGGGYQLN 362
                 |||  :       :    ::|: ::: |   |          : :  ::   :: ||:|:|| |  |
Db        305 SAQGPLSEELTVALLFRKGSPDTAIKYFFSIPLEEEVDYMYKFDNWIPCSVSCGKGVQTR 364

Qy        363 SAECVDIRLKRVVPDHYCHYYPENVKPKPKL-KECSMDPCPSSDGFKEIMPYDHFQPLPR 421
              :  |:| : |   | |  |      ||    ||:  | |      | :
Db        365 NLYCIDGKNKGRVEDDLCE---ENNATKPEFEKSCETVDCEA-----------------E 404

Qy        422 WEHNPWTACSVSCGG-GIQRRSFVCVEESMHGEILQVEEWKCMYAPKPKVMQTCNLFDCP 480
              |     | :|| :||  | | |     | :      :|  : ||:   |       :| |  |||| | ||
Db        405 WFTGDWESCSSTCGDQGQQYRVVYCHQVFANGRRVTVEDGNCT-VERPPVKQTCNRFACP 463

Qy        481 KWIAMEWSQCTVTCGRGLRYRVVLCINHRGEHVGGCNPQLKLHIKEECVIPIPCYKPKEK 540
              :| |   || |:   ||     :|| | |      | | |       : ||    : |
Db        464 EWQAGPWSACSEKCGDAFQYRSVTC---RSEKEG---EEGKLLAADAC------------ 505

Qy        541 SPVEAKLPWLKQAQELEETRIATEEP----TFIPEPWSACSTTCGPGVQVREVKCRVLLT 596
               | :          : :: :  |        |     ||:     |: | | |       : ||| |:
Db        506 -PAD-------EQEKFDTERTCNLGPCEGLTFVTGEWNLC-TRCNDTEETREVTCK---- 552

Qy        597 FTQTETELPEEECEGPKLPTERPCLLEACDESPASRELDIPLPEDSETTYDWEYAGFTPC 656
                           : :|    | |: ||::    | |              |         |:|  : :: |
Db        553 ----------DSQGRAYPLEK-CLVDNSTEIPTDTRSCATQP---PCEYEWTVSEWSKC 597

Qy        657 TATCLGGHQEAIAVCLHIQT--QQTVNDSLCDMVHRPPAMSQACNTEPCPPRWHVGSWGP 714
              |   |   ||:        :|      |        : |::    |        :|      :      | | |      |:   ||
Db        598 TTECGHGHKTRRVICAIHQNGGLEVVDEGHC-QAEKPEGKINCINEEKCIGIWYTSSWSE 656

Qy        715 CSATCGVGIQTRDVYCLHPGETPAPPEECRDE--KPHALQACNQFDCP 760
              |:| || | | |     ||:    : |   || | || ||    | ||   |||
Db        657 CTAECGGGSQDRVAVCINYDKKPV-PEWC-DEAVKPSEKQDCNVDDCP 702
```

RESULT 5
T00260
hypothetical protein KIAA0605 - human
C;Species: Homo sapiens (man)
C;Date: 01 Feb 1999 #sequence_revision 01 Feb 1999 #text_change 21 Jul 2000
C;Accession: T00260
R;Nagase, T.; Ishikawa, K.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, N.; Ohara,
DNA Res. 5, 31-39, 1998
A;Title: Prediction of the coding sequences of unidentified human genes. IX. The comp
A;Reference number: Z14386; MUID:98290545; PMID:9628581
A;Accession: T00260
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-951 <NAG>
A;Cross-references: EMBL:AB011177; NID:g3043733; PIDN:BAA25531.1; PID:g3043734
A;Experimental source: brain
C;Genetics:
A;Note: KIAA0605
C;Superfamily: thrombospondin type 1 repeat homology
F;46-106/Domain: thrombospondin type 1 repeat homology <THR1>

```
  Query Match             7.5%;  Score 700;  DB 2;  Length 951;
  Best Local Similarity   24.3%;  Pred. No. 1.1e-33;
  Matches  245;  Conservative  116;  Mismatches  354;  Indels  294;  Gaps   43;

Qy         32 AYFLPEFALSPQGSFLEDTTGEQFLTYRYDDQ-TSRNTRSDEDKDCNWDAWGDWSD---C 87
              |:||   |:                |:    |    |:  || :         |      |    ||:|:
Db         11 AWFLLVLAV---------VAGDTVSTGSIDNSPTSNSLEGGIDATAFW--WGEWTKWTAF 59

Qy         88 SRTCGGGASYSLRRCLT-----------GRNCEGQNIRYKTCSNHDCPPDAEDFRAQQCSA 137
              ||:|||| :    | ||                 | | |  : ||: |     :||||    || :|| :
Db         60 SRSCGGGVTSQERHCLQQRRKSVPGPGNRTCTGTSKRYQLCRVQECPPDGRSFREEQCVS 119

Qy        138 YNDVQYQGHYYEWLPRYND----PAAPCALKC-HAQGQNLVVELAPKVLDGTRCN-TDS 190
              :|    | |  ::| | | |           : || | |      ||       :|         ||| |   ||
Db        120 FNSHVYNGKIHQWKLE[illegible] 176
```

```
Qy        191 LDMCISGICQAVCDRQLGSNAKEDNCGVCAGDGSTCRLVRGQSKSHVSPEKREEN      246
                 :|:||  |: :|:|   |          | :|:|  |||:|         :|: :   : :
Db        177 RGVCVSGKCEPICSPVLEPCIHILDKCGIQGHASSS                 EHVTGNYRKGNAHG 228

Qy        247    VIAVPAGSRSVRIIVKGPAHLFIESK        ILQGSKGEHSFN         SPGVFVV 291
                  |  :|  |:|  :|           :|  |             |  :  ||        ||  :  :
Db        229 YSLVIHIPAGARDIQI            VERKKSADVLALADEAGYYIFNGNYKVDSPKNFNI 279

Qy        292 ENIIVEFQRGSERQIFKI        PGPLMADFIFKIRYIAAKDSVVQFFFY            336
                   : : :|                |           |   :   :   :
Db        280 AGIVVKYRRPMDVYETGIEYIVAQGPTNQGLNVMVWNQNGKSPSITFEYTLLQPPHESRP 339

Qy        337 QPISHQWRQT              GFFPCIVTCGGGYQLNSAECVDIRLKRVVPDHYCHYYP 384
              :||  :  :  :               | |     |   |   |:|        : |:      :|
Db        340 QPIYYGFSESAESQGLDGAGLMGFIPHN    GSLYGQASSE        RIGLDNRLFGHP 389

Qy        385 E NVKPKPKLKEPSMDP               PSSDGF                KEIMPYD HFQP 418
                :::  |    : : : |              |   ||                |:      |
Db        390 GLDMEGPSQGLEINEVCEQAGGACEGPPRGKGFRDRNVTGTPLTGDKDDEEVDTHFAS 449

Qy        419 LPRWEHNPWIACSVSCGG                     GIQRRS  -  FVCVEES         449
                 :  |  :  :  ,                        |  |  |      ||  |:
Db        450 QEFFSANAISDQLLGAGSDLKDFTLNETVNSIFAQGAPRSSLAESFFVDYEENEGAGPYL 509

Qy        450 MHGEILQVEEWK     CMYAPKPKV                        MQTCNLFDCPKWI 483
              ::|  |::    :          || | |                          |  :  :  :|
Db        510 LNGSYLELSSDRVANSSSEAPTPNVSTSLLTACNRIHKARIRPKARKQCVSPALHYRWK 569

Qy        484 AMEWSQCTVTCGRGLRYRVVLCINHRGEHVGG NPQLKLHIKEECVIPIPCYKPKEKSPV 543
                        |: ||  |:       :|: : |  |                       : |
Db        570 LSSHEPCSATCGIGVMSAYAMCVRYDGVEV     -   -     -   -  -  DDSYC 604

Qy        544 EAKLPWLKQAQELEETRIATE EPTFIPEPWSACSTTCGPGVQVREVKCRVLLTFTQTEI 602
              :|      |  :  :  :  |      | :|  :       | || ||| | | | |:: :|:       ::
Db        605 DA      LTRPEPVHEFCAGRECQPRWETSSWSECSRTCGEGYQFRVVRCWKMLSPGFDS 659

Qy        603 ELPEEECEGPKL   PTER  PCLLEACDESPASRELDIPLPEDSETTYDWEYAGFTPCTAT 659
                 :  : ||   :     | ||  |     ||                        |  :  : :  ||
Db        660 SVYSDLCEAAEAVRPEERKTCRNPACGP                          QWEMSEWSECTAK 700

Qy        660 CLGGHQEATAVCLHIQTQQTVNDSLCDMVHRPPAMSQACNTEPCPPRWHVGSWGPCSATC 719
              |        |    |    |:   :   :: |||   ||      : |      ||   :|  |   |||| :|
Db        701 C      GERCVVIRDTRCSE   DEKLCDPNTRPVG  EKNCTGPPCDRQWTVSDWGPCSGSC 753

Qy        720 GVGIQTRDVYCLHPGETPAPPEECRDE  KPHALQACNQFDCPPGWHIEEWQQCSRTCGGG 778
              | |     | |||                   |   :|:  | || |:  |     :||   |    ::|::: |||  |
Db        754 GQGRTIRHVYCKTSDGRVVPESQCQMETKPLATHPCGDKNCPAHWLAQDWER N TCGRC 813

Qy        779 TQNRRVTCRQLLTDGSFLNLSDELCQG  PKASSHKSCARTDCPPHLA---------VG 826
               : | | |                   ||   | |:   |      |        ||   :
Db        814 VKKRTVTCM                   ETANCRPQTRSCRECLARRPPRSTCEKRPCEKWYIS 860

Qy        827 DWSKCSVSCGVGIQRRKQVCQRLAAKGRRIPLSEMMCRDLPGFPLVRSC   875
                ||:|:   | ||:: !   | :                       |   :|| |
Db        861 PWSECTKTCGVGVRMRDVKCYQ                       GTDIVRGC 890


RESULT 6
I18517
procollagen N-endopeptidase (EC 3.4.24.14) I - bovine
N;Alternate names: procollagen N proteinase
C;Species: Bos primigenius taurus (cattle)
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: I18517
R;Colige, A.; Nusgens, B.V.; Lapiere, C.M.
submitted to the EMBL Data Library, February 1996
A;Description: Cloning of the cDNA of the bovine procollagen I N proteinase.
A;Reference number: Z18041
A;Accession: I18517
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1205 <COL>
A;Cross-references: EMBL:X96389; NID:g999769; PID:e228215; PIDN:CAA65251.1
```

```
A;Experimental source: skin
C;Genetics:
A;Gene: PCINP
C;Function:
A;Description: catalyzes cleavage of the propeptides of type I and II collagens prior
C;Keywords: hydrolase; metalloproteinase

  Query Match             6.4%;  Score 583.5;  DB 2;  Length 1205;
  Best Local Similarity   27.9%;  Pred. No. 3.3e-27;
  Matches  178;  Conservative   86;  Mismatches  266;  Indels  109;  Gaps   27;

Qy         74 KDGNWDAWGDWSDCSRTCGGGASYSLRRCLT        GRNCDGQNIRYKTCSNHDCPPDA 127
                :|||| ||   :   |||| | |   :   |:|          |     |      ::   :: |||
Db        554 KDGNWGAWSPWGSCSRTCGTGVKFRTRQCDNPHPANGGRTCSGLAYDFQLCNSQDCPDAL 613

Qy        128 EDFRAQQCSAYN         DVQYQGHYYEWLP  RYNDPAAPCALKCHAQGQNLVVELAPKV 180
               ||: :||  ::           | |          : |||  : |    | | | | ::    || :   |
Db        614 ADFREEQCRQWDLYFEHGDAQ        HHWLPHEHRDAKERCHLYCESKETGEVVSMKRMV 668

Qy        181 LDGTRCN TDSLDMCISGICQAVGCDRQLGSNAKEDNCGVCAGDGSTCRLVRGQSKSHVS 239
                |||||:   |:   :|:  | |: ||||   :| :  :||  |||| |, | | |:::|,          |
Db        669 HDGTRCSYKDAFSLCVRGDCRKVGCDGVIGSSKQEDKCGVCGGDNSHCKVVKGTFSR  S 726

Qy        240 PEKREE NVIAVPAGSRSVRIIVKGPAHLFIESKTLQGSK    GEHSFNSPG   VFVVEN 293
               |:|         :  :| |:|:| : |                    :  | |:   :      | :    |    |:
Db        727 PKKLGYIKMFEIPAGARHLLIQEADTTSHHLAVKNLETGKFILNEENDVDPNSKTFIAMG 786

Qy        294 ITVEFQRGSERQTFKIPGPLM    -                ADFIFKIRYIAAKDSV      VQFFFYQ 337
                 |::      |:! :  |||                      |           :|      :||:
Db        787 VEWEYRDEDGREFLQIMGPLHGTITVLVIPEGDARISLTYKYMIHEDSLNVDDNNVLEDD 846

Qy        338 PTGHQWRQTHFFFPCIVTCGGGYQLNSAECVDIRLKRVVPDHYCHYYPENVKPKPKLKEPC 396
                 :  ::|         : ||:   ||||  |           |                ::,        :,         ::|  |||       :  |
Db        847 SVGYEWALKKWSPCSKPCGGGSQFTKYGCRRRLDHKMVHRGFC       DSVSKPKATRRTC 902

Qy        397 SMDPCPSSDGFKEIMPYDHFQPLPRWEHNPWIACSVSCGG  GIQRRSFV CVEESMHGEIL 455
                :   |                              |  |      |     ||  ,||   |:|  ||   ||:       :
Db        903 NPQECSQ                             FVWVTGEWEPCSRSCGRTGMQVRSVRCVQPLHNNTTR 946

Qy        456 QVEEWKCMYAPKPKVMQTCNLFDCPKWIAMEWSQCTVTCGRGLRYRVVLCINHRGEHVG 514
                |       |   !  :|:    :  ||        || :| |    | |:|:|||| |   :   | |||            :   |
Db        947 SVHTKHCNDA RPEGRRACNRELCPGRWRAGSWSQCSVTCGNGTQERPVLCRTADDSEG 1004

Qy        515 GCNPQLKLHIKEECVIPIPCYK        PKEKSPVEAKLPWLKQA      QELEET       RIAT 563
                |    :   :              |  :    ||  :          |  :|| |         :   |   :                 :||        |
Db       1005 VCREE-RPETARICRLGPCPRNISDPSKKSYV      VQWLSRPDPNSPVQEISSKGRCQG 1059

Qy        564 EEPTF -- -- --     TPEPWSACSTTCGPGVQVREVKCRV                   LLTFTQTE 601
              : :   |                  ||          |   :|  |     :  :|    |                           |   |  :
Db       1060 DKSVFCRMEVLSRYCSIPGYNKLCCKSCNPHDNLTDVDDRAEPPSGKHNDIEELMPTLSV 1119

Qy        602 TELPEEECEGPKLPTERP     CLLEACDESPASRELDIP 637
                    |    |          |  :|  |   |                 |  ::   |   :       :|:|
Db       1120 PTLVMEVQPPPGIPLEVPLNTSSTNATEDHPETNAVDVP 1158


RESULT 7
T14764
hypothetical protein DKFZp434H204.1 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C;Accession: T14764
R;Wambutt, R.; Heubner, D.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, August 1999
A;Reference number: Z18181
A;Accession: T14764
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-898 <WAM>
A;Cross-references: EMBL:AL110229
A;Experimental source: adult testis; clone DKFZp434H204
C;Genetics:
A;Note: DKFZp434H204.1
```

```
  Query Match             6.3%;  Score 593;  DB 2;  Length 898;
  Best Local Similarity   24.3%;  Pred. No. 2.4e-27;
  Matches  269;  Conservative  111;  Mismatches  374;  Indels  354;  Gaps   58;

Qy        688 VHRPPAMSQACNTEPCPP---RWHVGSWGPCSATCGVGIQTRDVYCLHPGETPAPPEEC- 743
               :||       :|     :   ||       || | |   |: ||| |:| ::||||        |    | |
Db         29 IHR-----EAGGHDEVPPPVFSWHYGPWTKCTVTCGRGVQRQNVYCLERQAGPVDEEHCD 74

Qy        744 ----RDEKPHALQACNQFDCPPGWHIEEWQQCSRTCG-GGTQNRRVTCRQLLTDGSFLNL 798
                      |::      : |::  ||    |     |||   :|| ||    | | | :  :          |
Db         75 PLGRPDDQQ--RKCSEQPCPARWWAGEWQLCSSSCGPGGLSRRAVLCIRSVGLDEQSAL 131

Qy        799 SDELCQG-PKASSHKSCAR-TDCPPHLAVGDWSKCSVSCGVGIQRRKQVCQR---LAAKG 853
                    |:   |:   :      | |     ||     |||:||:|||:|| | |||    :|        :
Db        132 EPPACEHLPRPPTETPCNRHVPCPATWAVGNWSQCSVTCGEGTQRRNVLCTNDTGVPCDE 191

Qy        854 RRIPLSEMMCRDLPGFPLVRSCQMPECSKIKSEMKTKLGEQGPQILSVQRVYIQTREEKR 913
               :  |  ||:  |          ||    |: |                   ||   ||:
Db        192 AQQPASEVTC----SLPL---CRWP-----------LGTLGPE---------------- 216

Qy        914 INLTIGSRAYLLPNTSVIIKCPVRRFQKSLIQWEKDGRCLQNSKRLGITKSGSLKIHGLA 973
                       ||  ::  |  |  :   |                                             |  ||
Db        217 -GSGSGSSSHELFNEADFIP---------------------------------HHLA 239

Qy        974 APDIGVYRCIAGSAQETVVLKLIGTDNRLIARPALREPMREYPGMDHSEANSLGVTWHKM 1033
                          |      |: :             ||       |       |       |    :  ||        :
Db        240 P------RPSPASSPKP------GTMGNAIEEEA--PELDLPGPVFVD----------- 273

Qy       1034 RQMWNNKNDLYLDDDHISNQPFLRALLGHCSNSAGSTNSWELKNKQFEAAVKQGAYSMDT 1093
                            |   |   |   :  |:                  |     |  |  :      :|
Db        274               DFYYDYNFIN           FHEDLSYGPSEEPDL                      298

Qy       1094 AQFDELIRNMSQLMETGEVSDDLASQLIYQLVAELAKAQPTHMQ------WRGIQEETPP 1147
                            |    ||:                              :    |    |   :           :      ||
Db        299                ELAGTGD                   ETPPTHSRPAAPSTGSPVPATEPP 329

Qy       1148 AAQLRGETGSVSQSSHAKNSGKLTFKPKGPVLMRQSQPPSISFNKTINSRIGNTVYITKR 1207
                *|:    |    |    |  |        :|                   |     |||                         |
Db        330 AAKEEGVLGPWSPSPWPSQAG-------------PSPPPPSEQ              TPG 362

Qy       1208 TEVINILCDLITPSEA------TYTWTK---DG                TLLQPSV 1238
                   :||  |   :    ||    |              :   :|  :       ||                          :  |  |
Db        363 NPLINELPEEDTPIGAPDLGLPSLSWPRVSTDGLQTPATPESQNDFPVCKRCQSCLPPPW 422

Qy       1239 KIILSTTGKLQIQNPT***LIYEGVANGLGLVLSSSVIYALAIVIISVPENITEPP 1298
                :       |  |  ::       :       |  :|                 |  ||              || |        |   :    |        :  |
Db        423 R--DRTNEVFFPDEEPFGRG------APHL--PPRPSSTL      PPLSPVGSTHSSPSP 467

Qy       1299 NHLSVVVGGIV------PAALGANVTIRCPVKGV           PQPNITWLKRGGSLS-G 1342
                :    :    ||  |         |  ||     :        |    ||                  | |   :         |      ||    |
Db        468 DVAELWTGGTVAWEPALEGGLGPVDSELWPTVGVASLLPPPIAPLPEMK- VRDSSLEPG 525

Qy       1343 NVSLLF  NGSLLIQNVSLENEGTYVCIATNALGKAVATSVLHLLE  -   -RRWEL-R 1393
                     |              ||     ||  ::      ||::            ||                    |:  |               :    |||
Db        526 TPSFPAPGPGSWDLQTVAV--WGTFLPTTLTGLG--------HMPEPALNPGPKGQPESL 575

Qy       1394 IVELQGHKKYI-LQATNTRTNSND-PTGEP---------PPQPPF     WEPGNWSHC 1436
                        :        :  :        |  ::      ||:      |      :|                      ||   :|              |:  |||| |
Db        576 SPEVPLSSRLLSTPAWDSPANSHRVPETQPLAPSLAEAGPPADPLVVRNASWQAGNWSEC 635

Qy       1437 SATCGHLGARIQRP QCVMANGQEVSEALCDHLQKPLAGFE   CNIRDCPARWFTSVW 1492
                |  ||| |||  :  || :|             ::              |              :           :|          |::| |  |  |  |  :     |
Db        636 STTCG-LGA-VWRPVRCSSGRDED-----CAPAGRP----QPARRCHLRPC-ATWHSGNW 683

Qy       1493 SQCSVSCGRGYHSRQVTCKRTKANGTVQVVSPRACAP-KDRPLGRKPCFGHPCVQWEPGN 1551
                |:||  |||   |       |  |  '     |:         ::   :   |     |  |           :|        :||          ||:  |        :
Db        684 SKCSRSCGGGSSVRDVQCVDTR---DLRPLRPFHCQPGPAKPPAHRPCGAQPCLSWYTSS 740

Qy       1552   RCPGRCMGRAVRMQQRHTACQHNSSDSNCDDRKRPTLRRNCTSGACDVCWHTGPWKPC 1609
                   |       |  |              |||         |                    |::     ||         |  |    :      |          |      |||      |
Db        741 WRECSEACGGGE---QQRLVTC---PEPGLCEEALRPNTTRPCNTHPC-TQWVVGPWGQC 793
```

```
Qy       1610 TAACGRGFQSRKVDCIHTRSCKPVA-ERHCVQKKKPISWRHC-------LGPSCDRDCT 1660
               :| ||  |  |  |  |  |  |  |::|::   |         |   :   |  |  |  |                  |  |:||
Db        794 SAPCGGGVQRRLVKCVNTQTGLPEEDSDQCGHEAWPESSRPCGTEDCEPVEPPRCERDRL 853

Qy       1661 DTTHYCMFVKHLNLCSLDRYKQRCCQSC 1688
                 :   :|   ::  |    |  |     :  :||:||
Db        854 -SFGFCETLRLLGRCQLPTIRTQCCRSC 880
```

RESULT 8
T47158
hypothetical protein DKFZp762C1110.1 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C;Accession: T47158
R;Blum, H.; Bauersachs, S.; Mewes, H.W.; Weil, B.; Wiemann, S.
submitted to the Protein Sequence Database, March 2000
A;Reference number: Z24379
A;Accession: T47158
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-550 <AAA>
A;Cross-references: EMBL:AL162080
A;Experimental source: adult melanoma (MeWo cell line), clone DKFZp762C1110
C;Genetics:
A;Note: DKFZp762C1110.1

```
  Query Match             6.1%;  Score 571.5;  DB 2;  Length 550;
  Best Local Similarity   28.2%;  Pred. No. 2.3e-26;
  Matches  136;  Conservative   77;  Mismatches  164;  Indels  105;  Gaps   19;

Qy         75 GNWSAWSDWSDCSRTCGGGASYSLRRCLT      GRNCEGQNIRYKTCSNHDCPP DAE 128
                |:|   ||  |  |||||||||   |::|  |                |:  |||:  :||::|:   |||    :  :
Db        143 GSWGMWGPWGDCSRTCGGGVQYTMRECDNPVPKNGGKYCEGKRVRYRSCNLEDCPDNNGK 202

Qy        129 DFRAQQCSAYNDVQY    QGHYYEWLPRYN   DPAAPCALKCHAQGQNLVVELAPKVLD 182
                 ||  :||  |:|:                |       ||:|:|         |       |   |   |   |:|                |  |||:|
Db        203 TFREEQCEAHNEFSKASFGSGPAVEWIPKYAGVSPKDRCKLTCQAKGIGYFFVLQPKVVD 262

Qy        183 GTRCNTDSLDMCISGICQAVGCDRQLGSNAKEDNCGVCAGDGSTCRLVRGQSKSHVSPEK 242
                || |:  ||    :|:  | |      ||||  :   |     |  |  ||||  |:||||:  :  |        |    |  :
Db        263 GTPCSPDSTSVCVQGQCVKAGCDRIIDSKKKFDKCGVCGGNGSTCKKISG---SVTSAKP 319

Qy        243 REENVIAVPLGSRSVRITVKGPAHLFIESKTLQGSKGEHSF       NSPGVFVVEN      T 294
                        ::|  :|  |:  ::                          ::  :     :||:      ||             :  |  :::              :
Db        320 GYHKIIIIPAGATNIE                    VKHRNQKGSRNNGSFLAIKAADGTYILNGDYTLS 369

Qy        295 TVEFQRGSERQTEKTRGPLMADFTFK     TRYTAANDSVVQPPYQPISHQWRQTDEFPC 351
                |:|                                  |  ::|         ||:  :    ::  :          :  |:                     |
Db        370 TLE                       QDIMYKGVVLRYSGSSAALERIRSFSPLKE-------PL 404

Qy        352 TVTCGGGYQLNSAECVDIRLKRVVPDHYCHYYPENVKPKPKLKECSMDPCPSSDGFKEIM 411
                |:               |              :    ::|               |  ::                       |   |:   |    :    |:
Db        405 TIQV          LIVGNAESPKIK               YTYFV              KKKKESFNAIP                   437

Qy        412 PYDHFQPLPRWEHNPWTACSVSCGGGIQRRSFVCVEESMHGEILQVEEWKCMYAPKPKVM 471
                                   |          |     ||  ||     |  |||          |               ::|:                  :|         ||
Db        438               YFSAWVIEEWGECSKSCELGWQRRLVEC  RDINGQPAS          ECAKEVKPAST 484

Qy        472 QTCNLFDCPKWIAMEWSQCTVTCGRGLRYRVVLCINHRGEHVGGCNPQLKLHIKEEC 528
                :  |            ||:|       |||  |:  |||:|  :  |  :  |::|   |    :              |:|     |      |      :   |
Db        485 RPCADHPCPQWQLGEWSSCSKTCGKGYKKRSLKCLSHDGGVLSHESCDPLKKPKHFIDFC 544

Qy        529 VI 530
                 :
Db        545 TM 546
```

RESULT 9
T00017
gene ADAMTS-1 protein - mouse
C;Species: Mus musculus (house mouse)
C;Date: 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change 21-Jul-2000

```
C;Accession: I00017
R;Kuno, K.; Iizasa, H.; Ohno, S.; Matsushima, K.
Genomics 46, 466-471, 1997
A;Title: The exon/intron organization and chromosomal mapping of the mouse ADAMTS-1 gene
A;Reference number: Z14055; MUID:98110583; PMID:9441751
A;Accession: I00017
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-951 <KUN>
A;Cross-references: EMBL:AB0[illegible]; NID:g2809056; PIDN:BAA24501.1; PID:g2809057
A;Experimental source: strain 129SVJ
C;Genetics:
A;Gene: ADAMTS-1
A;Introns: [illegible]
C;Superfamily: thrombospondin type 1 repeat homology
F;542-598/Domain: thrombospondin type 1 repeat homology <THR>

  Query Match             [illegible]%;  Score 564.5;  DB 2;  Length 951;
  Best Local Similarity   27.7%;  Pred. No. 1.3e-25;
  Matches  133;  Conservative   72;  Mismatches  173;  Indels  103;  Gaps   18;

Qy         76 GNWDAWGDWSDCSRTCGGGASYSLRRCLT--  --GRNCEGQNIRYKICSNHDCPP-DAE 128
              |:|   || | |||||||||   |::| |              |: |||: :||::|:   |||      :
Db        544 GSWGEWGEWGDCSRTCGGGVQYTMRECDNPVPKNGGKYCEGKRVRYRSCNIEDCPDNNGK 603

Qy        129 DFRAQQCSAYNDVQYQGH---YYEWLPRYN--DPAAPCALKCHAQGQNLVVELAPKVLD 182
               || :|| |:|:                ||  |:|     |     |  |  | |:|          |  ||:|
Db        604 TFREEQCEAHNEFSKASFGNEPTVEWTPKYAGVSPKDRCKLTCEAKGIGYFFVLQPKVVD 663

Qy        183 GIRCNIDSLDMCISGICQAVGCDRQLGSNAKEDNCGVCAGDGSTCRLVRGQSKSHVSPEK 242
               || |: ||  :|: | |   |||| |  :  |  |   |||| |:|||||: :  |           |
Db        664 GTPCSPDSTSVCVQGQCVKAGCDRIIDSKKKFDKCGVCGGNGSTCKKMSG--TVISTRP 720

Qy        243 REENVIAVPLGSRSVRIIVKGPAHLFIESKTLQGSKGEHSP-    NCPGVPVVENTIVEF 298
                 ::: :| |: ::                 :: :   :||:    ||        : | : ::           |
Db        721 GYHDIVTIPAGAINIE               VKHRNQRGSRNNGSFLAIKAADGTYILNG    NF 767

Qy        299 QRGSERQTFKIPGPLMADFIFKTRYTAAKDSVVQFFFYQPISHQWRQIDFFPCTV-TCGG 357
                  :  |        | ::               ||:  :   : :  :    :  : |:                  | |:
Db        768 FLSTIEQDLIYKGTVL               RYSGSSAALERIRSPSPLKE              PLTIQVLMV 812

Qy        358 GYQLNSAECVDIRLKRVVPDHYCHYYPENVKPKPKLKECSMDPCPSSDGFKEIM--TYDH 415
               |   |                                      :|||:|                  |    :   :
Db        813 GHAL                                   -RPEIF                 EIYEMKKKIES 832

Qy        416 FQPLP    RWEHNPWTACSVSCGGGIQRRSFVCVEESMHGEILQVEEWKCMYAPKPKVMQ 472
               |   !     |        |    ''  :|  |  | |:|    |  :  :  |             :|     ||    :
Db        833 FNAIFIFSEWVIEEWGECSKTCGSGWQRRVVQCRDINGH        FASECAKEVKPASIR 886

Qy        473 TCNLFDCPKWIAMEWSQCTVTCGRGLRYRVVLCINHRGEHVG--GCNPQLK-LHIKEECV 529
                |       || |       :|| |: |||:| : | : |:| |  :          |:|   |    |   :  |
Db        887 PCAGGSTHWQVGDWSTCSKTCGKGYKKRTLKCVSHDGVLSNESCDPLKKPKHYIDFCT 946

Qy        530 I 530

Db        947 L 947


RESULT 16
T18856
metalloprotease inhibitor homolog [illegible] - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C;Accession: T18856; T24653
R;McMurray, A.
submitted to the EMBL Data Library, July 1995
A;Reference number: Z19031
A;Accession: T18856
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1444 <WIL>
A;Cross-references: EMBL:Z50004; PIDN:CAA90293.1; GSPDB:GN00028; CESP:C02B4.1
A;Experimental source: clone C02B4
R;McMurray, A.
submitted to the EMBL Data Library, July 1995
A;Reference number: Z19917
A;Accession: T24653
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1444 <WIL>
A;Cross-references: EMBL:Z50006; PIDN:CAA90302.1; GSPDB:GN00028; CESP:C02B4.1
A;Experimental source: clone F07C5
C;Genetics:
A;Gene: CESP:C02B4.1
A;Map position: X
A;Introns: [illegible]

  Query Match             5.2%;  Score 490;  DB 2;  Length 1444;
  Best Local Similarity   22.1%;  Pred. No. 6.1e-21;
  Matches  227;  Conservative   87;  Mismatches  299;  Indels  412;  Gaps   50;

Qy         75 DGNWDAWGDWSDCSRTCGGGASYSL     RRCLT         GRNCEGQNIRYKIC    S 119
              || |  ||          |  : |      | ::       | |:            |: ||| |||      |    |
Db        529 DGEWSEWGGAEKGCPTQQCAVSGSTIVQCQHRDCVNPAPNNGKICEGANIRCIVCGATS 588

Qy        120 NHDCPPDAEDFRAQQCSAYN-----DVQYQGHYYEWLPRYNDPAAPCALKCHAQGQNLV 173
              ::            |:|   : ||:             | |    |   :|                ||  :  ||    |    |:
Db        589 SNCLGFTREEFGNKTCSSTKYDPRKPDQQLIGEGFE        HSTQPCRVWCHLIGSELT 642

Qy        174 VELAPKVLDGIRCNTDSLDMCISGICQAVGCDRQLGSNAKEDNCGVCAGDGSTCRLVRGQ 233
                     :    |||  |   |:     |: |  |  |  |: ||  :               ||                |    :  |:
Db        643 RNKG-QFPDGIPCGFDA--YCVGGQCLALSCDNKALVEQPED--  ---        CPRIEGR 688

Qy        234 SKSHVSPE--KREENVIAVPLGSRSVRTIVKGPAHLFIESKTLQGSK---GEHSENSPG 287
              |    |      |           |  ::    ||  |  ||                   |  |      |  |        |  |        |
Db        689 S-VHQWEEWSSWSECSVSCGLGGREVR---            --ERKCSSGRKCQGVSEESRPCEG 736

Qy        288 VFVVENITVEFQRGSERQTFKTPGPLMADFIFKTRYTAAKDSVVQFFFYQPISHQWRQID 347
              |          |:                                                       :|:    :
Db        737 VL     RDCEEF       ----      -                       GEWK    E 749

Qy        348 FFPCTVTCGGGYQLNSAECVDIRLKRVVPDHYCHYYPENVKPKPKLKECSMDPCPSSDGF 407
               :  |:   |    |  |                   ||    |                          |    |  |  ||
Db        750 WGSCSEKCALGVQ      -              KRFRP                     CLTDQC SSKHL 778

Qy        408 KEIMTYDHFQPLPRWEHNPWTACSVSCGGGIQRRSFVCVEESMHGEILQVEEWKCMYAPK 467
               :|   |  |:            |:     |::||  |||||  : |       |:::      |: |: |
Db        779 QEERPCDNEGCWTNWDE--WSSCSQSCGGGRRYRTRKCLDDKCDGDDLEKE               827

Qy        468 PKVMQTCNLFDCPKWIAMEWSQCTVTCGRGLRYRVVLCINHRGEHVGGCNPQLKLHIKEE 527
                     :||      |        :|     |:|:||  |  :  |   ||        ||          | |  :      ::: 
Db        828 ---SCNTQKCTSQSWGDWLPCSVSCGTGFQTRERL          DGEDCATANKQARTCNQQ 879

Qy        528 CVIPTPCYKPKEKSPVEAKLP-WLKQAQELEETRIATEEPTFIPEPWSACSTTCGPGVQV 586
              |                     |       |     |   :     :                               |:  ||  |||  |:|
Db        880 C-      ---           PSAESLSVWSEWGE                          WTPCSATCGEGLQS 908

Qy        587 REVKCR---           -                      VLLTFTQTET               602
              ||    ||                                        |:::::   |
Db        909 RERSCRRGSCTEDDASQTRECVNGPCERSYLTWSEWTTCETCSSPLSRKRTAKCDGTTEN 968

Qy        603      ELPEEECE                  GPKLPTERPCLLEACDESPA SREEPTPLPPESETTYDW 648
                   :: || |:                         ||    |    ||  |:  :  :         |      |   ||  :|     :     |
Db        969 CQDKIDEETCTLACLREKHSPCPISP RRPKLTSNPCRKARCRPL    LPTESTHSEKW 1024

Qy        649 EYAGFTPCTATCGGG              QEATAVCLHTQTQQTVNESGSCHMVHRPLAMSQACNTE 701
                    |      ||: :|   |                   |||             ||||::                                  ||
Db       1025 SEWG   PCSVTCGSGRRVRTRGCQEASCPEQHTQTEE                          CNLN 1063

Qy        702 PCPPRWHVGSWGPCSATCGV-GIQTRDVYCLHPG      ETPAPPEECRD                745
                  |     :         |      ||  :||   |||||      ||                      :  |          |:|
Db       1064 [illegible] 1123

Qy        746          -EKPH-        ALQACN                                            CTDCPP           761
                           :  |                :   ||:                                              |    |   |
Db       1124 TSENGFDAPRWSEWSSWSACSCESLTSTRRREQVVPPIVQCTCAGATLEQTPCAPGSCS 118[illegible]
```

```
Qy      762 --GWRI-EEWQQCSRTCGGGTQNRRVTCRQLLTD--GSFLN          LSDELCQ 804
               || :  ||  ||: ||  | | |    | : :      |:: :           : | :|
Db     1184 PSAGGWSLWSEWSSCSKTCGDTCHQIRNRMCSEPIPSNRGAYCSGYSPDQRPCVMENVCS 1243

Qy      805 GPKA-------------------SSHKSCAPTDCPP------------------------- 821
               |                   : |:|   |    |
Db     1244 DEKVPCGWTPWTAWSECTIDYCRNGHRSRTRPCANPKPSQGGAQCTGSHTTLNPCEDFAEC 1303

Qy      822 HLAVG      -DWSKCSVSCGVCIQRRKQVTCSLAAFGRR---GEMMCRHLPCFIVRSC 875
            ||   |     ||: || |||  |:| | : |      || :         |  |        |
Db     1304 HLRDGGWSTWSDWTPCSASCGGGVQTRDRSCSSPRPKGGQ------SCSGLA--HQTSLC 1355

Qy      876 QMPEC 880
             :| |
Db     1356 DLPAC 1360


RESULT 11
T09355
hypothetical protein KIAA0688 - human
C;Species: Homo sapiens (man)
C;Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 21-Jul-2000
C;Accession: T00355
R;Ishikawa, K.; Nagase, T.; Suyama, M.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, N.
DNA Res. 5, 169-176, 1998
A;Title: Prediction of the coding sequences of unidentified human genes. X. The complete
A;Reference number: Z14142; MUID:98403889; PMID:9734811
A;Accession: T00355
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-837 <ISH>
A;Cross-references: EMBL:AB014588; NID:g3327189; PIDN:BAA31663.1; PID:g3327190
A;Experimental source: brain
C;Genetics:
A;Gene: KIAA0688
C;Superfamily: thrombospondin type 1 repeat homology
F;519-575/Domain: thrombospondin type 1 repeat homology <TBR3>

  Query Match             4.1%;  Score 415;  DB 2;  Length 837;
  Best Local Similarity   33.8%;  Pred. No. 8.6e-17;
  Matches   97;  Conservative   40;  Mismatches  122;  Indels   28;  Gaps   11;

Qy       76 CNWDAWCDWSDCSRTCGGGASYSLRRCLT       CRNCEGQNIRYKTCSNHDCPP-DAE 128
            | |  || | ||||||||   :| | |        : |||:  |:::|:  |||   |
Db      521 GGWGPWGPWGDCSRTCGGGVQFSSRDCTRPVPRNGGKYCEGRRTRFRSCNTEDCPTGSAL 580

Qy      129 DFRAQQCSAYND-     VQYQGHYYEWLPRYND  PAAPCALKCHAQGQNLVVELAPKVL 181
            || :||:|||          : |    :|:|||           | | | |:        | |:|:
Db      581 TFREEQCAAYNHRTDLFKSFPGP-MDWVPRYTGVAPQDQCKLTCQARALGYYYVLEPRVV 639

Qy      182 DGTRCNTDSLDMCISGICQAVGCDRQLGSNAKEDNCGVCAGDGSTCRLVRGQSKSHVSPE 241
            ||| |: ||  :|: | |   |||| :||  | | | || |||| |       || |
Db      640 DGTPCSPDSSSVCVQGRCIHAGCDRIIGSKKKFDKCMVCGGDGSGC---SKQSGSFRKFR 696

Qy      242 KREENVIAVPLGSRSVRITVKG PAH  LFIESKTLQGS   KGEHSFNSPGVFVVENTT 295
                 ||: :| |:  : :  :| | |  :::  |   ||    ||::         ||
Db      697 YGYNNVVTIPAGATHILVRQQGNPGHRSIYLALKLPDGSYALNGEYTLMPSPTDVVLPGA 756

Qy      296 VEFQRG   SERQTFKIPGPLMADFIFKTRYTA AKDSVVQFFFYQP 328
            |  :       :  :|    |||        :          :|: ::: |: |
Db      757 VSLRYSGATAASETLSGHGPLAQPLTLQVLVAGNPQDTRLRYSFFVP 803


RESULT 12
T15976
hypothetical protein F08C6.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C;Accession: T15976
R;Bentley, D.
submitted to the EMBL Data Library, June 1995
A;Description: The sequence of C. elegans cosmid F08C6.
A;Reference number: Z18440
A;Accession: T15976
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-957 <BEN>
A;Cross-references: EMBL:U29378; NID:g868184; PID:g868185; PIDN:AAA68721.1; CESP:F08C
A;Experimental source: strain Bristol N2
C;Genetics:
A;Gene: CESP:F08C6.1
A;Introns: [illegible]

  Query Match             3.8%;  Score 382;  DB 2;  Length 957;
  Best Local Similarity   20.5%;  Pred. No. 5.8e-13;
  Matches  169;  Conservative   61;  Mismatches  216;  Indels  380;  Gaps   39;

Qy       75 DCNWDAWCDWSDCSR-TCGG-----CASYSLRRCLT----CRNCEGQNIRYKTCSNH- 121
            ||:|   ||   : ||: ||  |         | : :  |  |               | :| |     |     ||
Db      431 DGSWSGWGA-TICSQCICNGILGSVGLAIARRTCSAPYPANGGSDCVGSISRAVLCSRQC 489

Qy      122 -DCPPDAEDFRAQQCSAY----NDVQYQGHYYEWLPRYNDPAAPCALKCHAQ-      - 168
                   ::: : :|         || :  |    : | |:  |    | : | |
Db      490 CPASKAVSCYISKKCMSQRRTKNCWELLDPRCLQ-LNPE  CQEACRVRCSVQQHTCSQRN 546

Qy      169 ----GQNLVVELAPKVLDGTRCNTDSLDMCISGICQAVGCDRQLGSNAKEDNCGVCAGDG 224
                 | ||              ||| |   |      |: | | |: |:       :||           |   |
Db      547 YRPPGDNLP         DGTSGGYDR   YCLIGECLALNGN     NNALISRDQSCPTD- 591

Qy      225 STCRLVRGQSKSHVSPEKREENVIAVPLGSRSVRITVKGPAHLFIESKTLQGSKGEHSFN 284
             ||                         |:   :|                       |    :|   |   |
Db      592  TC            -          -PITPQC                      GCVYRGQWGTWGL 612

Qy      285 SPGVFVVENTTVEFQRGSERQTFKIPGPLMADFIFKTRYTAAKDSVVQFFFYQPISHQWR 344
                                                                          |
Db      613 -------------------------- -----     ---      ---       --    WT 614

Qy      345 QTDFFPCTVTCGGGYQLNSAECVDIRLKRVVPDHYCHYYPENVKPKPKLKECSMDPCPSS 404
                  || ||||||:                || :  |      :       : : : || : |||
Db      615 S--    -CTATCGGGYR-------------KR---NRACSITGQCEGNEDEIEVCSSESCPS- 654

Qy      405 DGFKEIMPYDHFQPLPRWEHNPWTACSVSCGGGIQRRSFVCVEESMHGEILQVEEWKCMY 464
                   :  :  :     | |        |||||| |  |  |     |: |  |  :
Db      655       VLRVGNEWSTWTEWNH        CSVSCGRGLQARYRKCL    DFRRTL---      - 693

Qy      465 APKPKVMQTCNLFDCPK                WIAMEWSQCTVTCGRGLRYRVVLCINHRGEH 512
                       ||||                   :  |   :  |:|  |  :   |        |: :
Db      694 -----------AFDCPGENKVINELRITTPKARSYIMGSVRCNKIKEN      TISEKNIE 738

Qy      513 VGGCNPQLKLHIKEECVIPIPCYKPKEKSPVEAKLPWLKQAQELEETRIATEEPTFIPEP 572
            |  |                           :  |   |    |
Db      739 VRSC-------------------- ---DNGPCNAIGVWGTWG-----      ---  ---G 758

Qy      573 WSACSTTCGPGVQVREVKCRVLLTFTQTETELPEEECEGPKLPTERPCLLEACDESPASR 632
            || |||:||||  ||                                   :|  |  | || |    |
Db      759 WSTCSTSCGPGTLVR---------------------------     ---QRTCNREPCDGSAHER 789

Qy      633 ELDIPLPEDSETTYDWEYAGFTPCTATCLGGHQEAIAVCLHIQTQQTVNDSLCDMVHRP 691
                                     |   |||                        || :   :
Db      790 R-------------- ------ SCNVATC                        QNDSIWSL      805

Qy      692 PAMSQACNTEPCPPRWHVGSWGPCSATCGVGIQTRDVYCLHPGETPAPPEECRDEKPHAL 751
                                |:    |   ||  ||  |:::|     |      |     |   |:
Db      806                     WN   EWSPCSRVCGRGLRSRSRGCPGSGCMGASSEQ---  ---- 839

Qy      752 QACNQFDCPPG-------WRIEEWQQCSRTCGGGTQNRRVTCRQLLTDGSFLNLSDELCQ 804
             | ||:   |            |      | ||| :|| | : |  |||        |::
Db      840 QFCNEQACASSSANDWGTW--SGWSQCSVSCGACVKRRTRTCRTGNCPGNY---------- 888

Qy      805 GPKASSHKSCARTDCPPHLAV----CDWSKCSVSCGVCIQRRKQVC 846
               | |:    |   ||     |        | || || :|| |:::| : |
Db      889 --KESA   ICNDRDCENKNAAWGGWGYWSSCSETCGDGVRKRVRKC 930
```

RESULT 13

```
T25061
hypothetical protein T21B6.3 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C;Accession: T25061
R;Cottage, A.
submitted to the EMBL Data Library, November 1995
A;Reference number: Z19975
A;Accession: T25061
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-788 <WIL>
A;Cross-references: EMBL:Z68011; PIDN:CAA92014.1; GSPDB:GN00028; CESP:T21B6.3
A;Experimental source: clone T21B6
C;Genetics:
A;Gene: CESP:T21B6.3
A;Map position: X
A;Introns: 20/1; [illegible]

  Query Match             3.7%;  Score 341.5;  DB 2;  Length 788;
  Best Local Similarity   21.7%;  Pred. No. 1.9e-12;
  Matches  192;  Conservative   78;  Mismatches  304;  Indels  311;  Gaps   47;

Qy         62 DQTSRNIP  [illegible]  119

Db        144 DKVDRKKRQQGDEAVEGVGD WIDWSHCSSN      CHEVF     SQACEYGRKIQRRGC 193

Qy        120 GNHDCP       PDAEDFRAQQCSAYNDVQYQCHYYEWLPRYNDPAAPCA LKCHAQGQNL 172

Db        194 PARSAPQRVPAPPAQQY APRAPEYPSAQQQQQQREQQQREQQHREHQARLQQHQQQQQQ 252

Qy        173 VVE                           LAPKVLDGIRCNIDSLDMCISGICQAVGCDRQLGS 210

Db        253 QQQQQQRPPQPQPQPQPQPPQRPPQQPQSPSGI                  HELHGQRQ     294

Qy        211 NAKEDNCGVCAGDGSTCRLVRGQSKSHVSPEKREENVIAVPLGSRSVRITVKGPAHLFIE 270

Db        295   REQQ                    QQQQQQQQQQQRQQNPQQQP  -                319

Qy        271 SKTLQGSKGEHSFNSPGVFVVENTTVEFQRGSEPQTFKIPGPLMADFIFKTTYTAAKDSV 330

Db        320 QQTIQPGQSQTQLQSGPVPPQQHPQ    QQQPQQQPELERSPC                 LQH 362

Qy        331 VQFFFYQPISHQWRQT         DFFPCTVTCGGGYQLNSAECVDI      RLKRVVPDHYC 380

Db        363 AQ  LYQQRMGQYRENPNQPHPARPKAPSCSNQP       QAEVPQAPQQPPRQPVTA 414

Qy        381                     HYYPENVKPKPKLKECSMDPCPSSDGPKEIMPYDHFQPLPR  W EHN 425

Db        415 PVINIAIQPPLPQPYPPRYKPAP          PPPPACDGQGCVNP     PVVSGVWHDWS 462

Qy        426 PWIACSVSCGGGIQ[illegible]TENWCA 464

Db        463 DWSIPSCIPGDGAKGRKKPCGINNCQGAPYETE                PNIGPCQIWSEW 508

Qy        465 MEWSQCTVTCGRGLRYRVVLCINHRGEHVGGCNPQLKLHIKEECVIPIPCYKPKEKSPVE 544

Db        509 CEWSTCSASCGSGQRERIRPC         HEGTNRCEGKDYESRQC             SACPCE 552

Qy        545 AKLPWLKQAQELEEIRIATEEPTFIPEPWSACSITCGPGVQVREVKCRVLLTFTQTETEL 604

Db        553 EWSQW                          EPWGQCSVTCGQGVAVRQRIC             LNGVE 580

Qy        605 PEEECEGPKLPTERPCLLEACDESPASRELDIPLPEDSETTYDWEYAGFTPCTATCLGG 663

Db        581 GDHLCQGPK   TEQ      RACDGGPCSL                   WSPWQEWSTCSASCGSGC 622

Qy        664 HQEAIAVCLHIQTQQTVNDSLCDMVHRPPAMSQACNTEPCPPRWHVGSWGPCSATCGVGI 723

Db        623 MKRRQRVCQPGTDCQGPNEE                 SQPCYGPPCAEWIEWCEWSQCSSKCGPGQ 671

Qy        724 QTRDVYCLHPGETPA       PPEE    CRDEKPHALQACNQFDCPPGWHIEEWQQCSRSC 775
```

```
Db        672 RTRTRGCLGPNGQEATTCQGPSTETTDCEGQ      SCCNWSE      WC  EWSMCDKEC 720

Qy        776 GGGTQNRKVTCRQLLTDGSFLNLSDELCQGPKASSHKSCARTDCPPHLAVGDWSKCS   V 833

Db        721 GGG QVRYT                                   EYMFRIGC          EWSPCSIQL 745

Qy        834 SCGVGIQRRKQVCQRLAAKGRRIPLSE  MMCRDLPGFPLVRSC 875

Db        746 ACEVGVQSRSRQC  VGESGCHCTGLAEESQQCRGLIQCPPKEPC 788


RESULT 14
T00028
brain specific angiogenesis inhibitor 3 - human
N;Alternate names: BAI3 protein
C;Species: Homo sapiens (man)
C;Date: 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change 21-Jul-2000
C;Accession: T00028
R;Shiratsuchi, T.; Nishimori, H.; Ichise, H.; Nakamura, Y.; Tokino, T.
Cytogenet. Cell Genet. 79, 103-108, 1997
A;Title: Cloning and characterization of BAI2 and BAI3, novel genes homologous to bra
A;Reference number: Z14666; MUID:98194217; PMID:[illegible]
A;Accession: T00028
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1522 <SHI>
A;Cross-references: EMBL:AB005299; NID:g3021700; PIDN:BAA25063.1; PID:g3021701
A;Experimental source: brain
C;Genetics:
A;Gene: GDB:BAI3
A;Cross-references: GDB:9838096; OMIM:602684
A;Map position: 6q12-6q12
C;Superfamily: thrombospondin type 1 repeat homology
F;344-398/Domain: thrombospondin type 1 repeat homology <TBR3>

  Query Match             3.1%;  Score 292.5;  DB 2;  Length 1522;
  Best Local Similarity   19.5%;  Pred. No. 3.8e-09;
  Matches  233;  Conservative  134;  Mismatches  397;  Indels  433;  Gaps   62;

Qy        327 FIFKT          TYTAAKDSVVQFFFYQPISHQWRQTDFFP        CTVTCGG          357

Db         10 YIESIYLLVMPGFNAAQDEWCSILVKGVIYGSYSVSEMFPKNFINCTWTLENPDPTKYSI 69

Qy        358    GYQLNSAECVDIRLKRVVPDHYCHYYPENVKPKPKLKECSMDPCPSSDGF            407

Db         70 YLKTTKKPLCTTNFSLLAYQIPHHCH    ERTRGLLPKNRGTMQLTNCTENZIVELQYPEN 126

Qy        408      KEIMP                      YDHFQPLPRWEHNPWTACSVSCGGGIQRRSFVCVE 447

Db        127 FIQIRRVPPINFPGLQKKGEEDQKSPFEEFLVLNKVSPSQF GCHVLC             IWLE 176

Qy        448 EAMRGLTLQEVLLWKCMYAPKTKVMQGCNGTCPKWIAMEW    SQCTVTCGRGLRYRVVL 504

Db        177 SCLKSENGRTESCGIMYT        KCT     CPQHDG EWGIDDQSLI        LLNNVVL 220

Qy        505 CINHRGEHVGGCNPQLKLHIKEECVIPIPCYKPKEKS          PVEAKLPWLKQAQELE 557

Db        221 PLNEQTE     GCLTQ ELQTTQVCNLTREAKRPPKEEIGMMGDHTTKSQRP    RSVH 271

Qy        558 ETRIATEE---PTFIPE-----------PWSACSITCGPGVQVREVKCRVLLTFTQTETEL 604

Db        272 EKRVPQEQADAAKTMAQTGESGVLEWSQWSTCSVTCGQGSQVR                    314

Qy        605 PEEECEGPKLPTERPCLLEACDESPASRELDIPLPEDSETTYDWEYAGFTPCTATCLGGH 664

Db        315  -    -          TRTCV        SPYGTHCSGPLRE                    332

Qy        665 QEATAVCLHIQTQQTVNDSLCDMVHRPPAMSQAC NTEPCPPR  WHVGS WGPCSATCG 720

Db        333      -                              SRVCNNTALCPVHGVWEEWSPWSQCSFTCG 362

Qy        721 VGIQTRDVYCLHPGETPAPPEECRDEKPHALQACNQFDCPPGWHIEE    WQQCSRTCGG 777
```

```
Db        363 RGQRTRTRSCTPPQYGGRP    CEGPETHH-KPCNIALCPVDGQWQEWSSWSQCSVTCSN 418

Qy        778 GIQNRPVICRQIHTDGSFINLSDPICQGPKASSHKSCARTDCPPH---LAVCDWSKCSVS 834
              |||  |     |          ||           |:||  |  |  :  |    :|   :        |  ||  ||  |
Db        419 GTQQRSRQCIAAAHCCS-------FCRCPWAPS-RPCYNPPCTANCQWNQWCHWSQCSKS 470

Qy        835 CGVGIQRRKQVCQRLAAKGRRIPLSEMMCRDLPGFPLVRSCQMPECSKIKSEMKTKLGEQ 894
              |   |  :||  :  ||        ::                          :|      |::     ::
Db        471 CDGGWERRIRTCQGAVITGQ-----------------------QCEGTGEEVRRCSEQR 506

Qy        895 GPQILSV-QRVYIQTREEKRINLTIGSRAYLLPNISVIIKCPVRRFQKSLIQWEKDGRCL 953
               |        :       |:  :     ||       |   |:                  :||:         :            ||
Db        507 CPAPYEICPEDYLMSMVWKR--TPAGDLAF--------NQCPLNAIGTT------SRRC- 549

Qy        954 QNSKRLCIIKSCSIKIHGLA    AP-IGVTWPIACSAQEIVVLKLIGIDNRLIARPALEE 1010
                                   || :|||:|        |        |||:
Db        550 --------------SLSLHGVAFWEQPSFA--RCIS------------------------- 569

Qy       1011 PMREYPGMDHSEANSLGVTWHKMRQMWNNKNDLYLDDDHISNQPFLRALLGHCSNSAGST 1070
                 ||   :  ||      |                       |        |    |  :|
Db        570 --NEYRHLQHSIKEHLA------------KGQRMLAGDGMS------------------ 596

Qy       1071 NSWELKNKQFEAAVKQGAYSMDTAQFDELIRNMSQLMETGF---VSDDLASQLIYQLVAE 1127
                 ::          :      ::    |:  |          |::||::         :                 ||  :    |    :|:|:
Db        597 ---QVTKTLLDLTQRKNFYAGDLLMSVEIIRNVTDTFKRASYIPASDGV--QNFFQIVSN 651

Qy       1128 LAKAQPTHMQWRGIQEETPPAAQLRGETGSVSQSSHAKNSGKLTFKPKGPVLMRQSQPPS 1187
              |    :         :|      |:    |  :  :|            :          |         |  :  |:
Db        652 LLD-EENKEKWEDAQQIYPGSIEL---MQVIEDFIHIVGMGMMDFQ-------------- 693

Qy       1188 ISFNKTINSRI--GNTVYITKRTEVINILCDLITP---SEATYIWIKDGTLLQPSVKIIL 1242
                          ||  :   || |     ::        ::| |:   |         :        |  ::               ::::
Db        694 -------NSYLMTGNVVASIQKLPAASVLTDINFPMKGRKGMVDWARNS-----EDRVVI 741

Qy       1243 DGTGKIQIQNPIRKEQGIYECSVANHLGSDVESSSVLYAEAPVILSVERNITKPEHNHLS 1302
                  :      :  :         |  ::                 ||            :|||         :||       || |
Db        742 PKSIFTPVSSKELLDESSVFV------LG------AVLYKNLDLILPTLRNYT-------- 781

Qy       1303 VVVGGIVEAALGANVTIRCPVKGVPQPNITWLKRGGSLSGNVSLLFNGSL----LLQNVS 1358
              |:     |:                 ||:|           |:|    |                |      ::  |   ||:|          .|   .   |
Db        782 VINSKIIV------VTIR------PEPKTT----DSPLEIELAHLANCTLNPYCVLWDDS 825

Qy       1359 LENE--GTYVCIATNALGKAVATSVLH---LLERRWPESRIVFLQGHKKYILQATNT 1410
                  ||   ||:    :|      |  |  .     |      |  :|        :    :    |    ::  |::::  |
Db        826 KTNESLGIW---SIQSR-KIVLIDASHINCLCDRL- SIFAILAQQPREIIMESSGT 876
```

```
RESULT 15
T00026
brain-specific angiogenesis inhibitor 1 - human
N;Alternate names: BAI1 protein
C;Species: Homo sapiens (man)
C;Date: 22-Jan-1999 #sequence_revision 22 Jan 1999 #text_change 12-Feb-1999
C;Accession: T00026
R;Nishimori, H.; Shiratsuchi, T.; Urano, T.; Kimura, Y.; Kiyono, K.; Tatsumi, K.; Yoshid
submitted to the EMBL Data Library, June 1997
A;Reference number: Z14064
A;Accession: T00026
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1584 <NIS>
A;Cross-references: EMBL:AB005297; NID:d1175073; PID:d1024528
A;Experimental source: brain
C;Genetics:
A;Gene: GDB:BAI1
A;Cross-references: GDB:9838098; OMIM:602682
A;Map position: 8q24-8q24
C;Superfamily: thrombospondin type 1 repeat homology
F;408-462/Domain: thrombospondin type 1 repeat homology <THR3>

  Query Match             3.1%;  Score 289.5;  DB 2,  Length 1584;
  Best Local Similarity   20.8%;  Pred. No. 6e-09;
  Matches  230;  Conservative  138;  Mismatches  388;  Indels  351;  Gaps   60;
```

```
Qy        487 WSQCTVTCGRGLRYRVVLCINHRGFHVGGCNPQL---KLHIKEECVIPIPCYKPKEKSPV 543
              |  :||   ||  ||:  |   |:    |   |||    |     :    :|  |                |   ::
Db        270 WGECTRDCGGGLQTRTRTCLPAPGVEGGGCEGVLEEGRQCNREAC       GPACRTSS 322

Qy        544 EAKLPWLKQAQELEETRIATEEPTF-IPE------     --PWSACSITCGPGVQVREVKCR 592
                ::          |:   ||           ::    |    |:                  |||  ||:|||  |  |  |     |
Db        323 RSQSLRSTDARRREELGDELQQFGFPAPQTGDPAAEEWSPWSVCSSTCGEGWQTRTRFCV 382

Qy        593 VLLTFTQTETELPEEECEGPKLPTERPCLLEA-CDESPASRELDIPLPEDSETTYDW-EY 650
                     ||             |  ||  |    :|  |      |  |          |                      |  |:
Db        383 SSSYSTQ---------CSGP-LREQRLCNNSAVCPVHGA-----  ---  --  ----WDEW 415

Qy        651 AGFTPCTATCLGGHQEATAVCLHIQTQQTVNDSLCDMVHRPPAMSQACNTEPCPPR---  706
              :  ::  |::||   |  ::         |      |                   :  |:         |      ::  ||       || |
Db        416 SPWSLCSSTCGRGFRDRTRTCRPPQ----FGGNPCE---GPEKQIKFCNIALCPGRAVDG 468

Qy        707 -WHV-GSWGPCSATCGVGIQTRDVYCLHPGETPAPPEECRDEKPHALQA--CNQFDCP-- 760
               |.     ||    |||.|    |   |   |       |     |       |       ||:        |  ::        |          ||
Db        469 NWNEWSSWSACSASCSQGRQQRTRECNGPSYGGA---ECQG---HWVETRDCFLQQCPVD 522

Qy        761 ---PGWHLPEWQQCSRTCGGGTQNRRVTCRQLLTDGSFLNLSDELCQGPKASSHKSCART 817
                        |       |    ||  |||  |:|  |          |                 |  |           ||||:      ::  |
Db        523 GKWQAW--ASWGSCSVTCGAGSQRRERVC-----SGPF--FGGAACQGPQ-DEYRQCGTQ 572

Qy        818 DCP-PHLAVGDWSKCSVSCGVGIQRRKQVCQR---LAAKGRRIPLSEMMC----RDLPGF 869
               || ||                                      ::|         |     :     |    |:               |:   |
Db        573 RCPEPH--------------------EICDEDNFGAVIWKETPAGEVAAVRCPRNAIGL 611

Qy        870 PLVRSCQMPECSKIKSEMKTKLGFQGPQILSV--QRVYIQTREEKRINLTIGSRAYLLPN 927
                ::|  |::  |              |    |   :          :  :|:   :   :   :  |||          :|         |
Db        612 -ILRRCELDEEGIAYWEPPTYI----RCVSIDYRNIQMMTRE---HLAKAQRGLPGEG 661

Qy        928 TSVIIKCPVRRFQKSLIQWEKDGRCLQNSKRLGITKSGSLKIHGLAAPDI---------- 977
                |  :|                ::|::    :|                   |  :  ||  |          |:      |:
Db        662 VSFVI--------QTLVFISQD----------GTSYSGDL----LSTIDVLRNMTEIFRR 699

Qy        978 GVYRCIAGSAQETVVL--KLIGTDNR------LIARPALREPMREYPGMDHSEANSLGVT 1029
                 |       |     |      |    :           |:      :||                  :|   |       :!     |                         :   :   |
Db        700 AYYSFTPGTVQNFVQILSNLLAFNRTPWL-AQLAGLNARLLR        LVLDFVDVI 752

Qy       1030 WHKMRQMWNNKNDLYLDDDHISNQPFLRALLGHCSNSAGSTN          SWR  -  IKNK 1078
                :|:   :                |     |        |::                |         |        ::|:|:                         |                 |
Db        753 GFRMKDL----RDAYQVTDNL-      ----VLSIHKLPASGATDISFPMKGWRATGDWAKVP 801

Qy       1079 QFEAAVKQCAYSMDTAQFDE              LIENMSQLMETGEVSDDLASQLIYQLVAELA 1129
                 :        |   :    :|             :   ||               | ||:       :            :     |    | :::        |
Db        802 EDRVTVSKSVFSTGLTEADEASVIPVVGTVLYKNLGDFLALQRNLTVLNSKVISVTVKPPP 861

Qy       1130 KAQPT-          HM                                 QWRGIQEETPPAAQLR  1152
              ::     |          ||                                 |||   :      |    |        ||
Db        862 RSLRTPLETEFAHMYNGTTNQTCILWDETDVPSSSAPPQLGPWSWRGCR--IVPLDALRI 919

Qy       1153 ------ GETGSVSQSSHAKNSGKLTFKPKGPVLMR------------ ---   -----QS 1183
                            ::|   |      |      |   |      |         :::                                        :|
Db        920 RCLCDRLSTFAILAQLSADANMEKAIL PSVILIVGGGVSSLTILMLVIIYVSVWRYIRS 978

Qy       1184 QPPSISFNKTINSRIGNTVYITKRTEVIN-ILCDLITP-----SEATYTWTKDGTLLQPS 1237
               :         |      |       ::             |   :   :     :|:     |    ::|  |  :                 :::  |                :|       :
Db        979 ERSVILINFCLSIISSNALILIGQTQTRNKVMCTLVAAFLHFFFLSSFCW- -- VLTEA 1033

Qy       1238 VKIILDGTGKIQ                            -     -IQNPTKKEQGIYECVANHLGSVESS 1276
                :        :        ||   ::                                               |           |   :|            |      |:              :|
Db       1034 WQSYMAVTGHLRNRLIRKRFLCLGWGLPALVVAISVGFTKAKGY---STMNYCWLSLE G 1089

Qy       1277 SVLYA-     -LAIVILSVLRNIIKILHNHLSVVVG              GTVLAAL        GANV 1317
                :|||               |    |:|                          :::|:|                 ||    :     |          ||::
Db       1090 GLIYAFVGPAAAVVL                    VNMVIGILVFNKLVSKDGIIDKKLKERAGASL 1136

Qy       1318 TIRCPVKGVPQPNITWLKRGGSLSGNVSLLFNGSLLLQNVSLENEGTYV-    -CIATNAL 1373
                      |   |     :|         :||:         :::          |   ||          :|      |         ||      :          ||            :
Db       1137 WSSCVV--LPLLALTWMSAVLAVTDRRSALFQ---ILFAVFDSLEGFVIVMVHCILRREV 1191
```

```
Qy       1374 GKAVAISVLHLLERRWPFSRIVFIQGH 1400
                 :|     |:    |       .:    |   |:
Db       1192 QDAVKCRVVDRQEEPNCESGISFQNGH 1218

Search completed: March 22, 2003, 13:31:49
Total time : 59 secs
```

```
                    GenCore version 5.1.4_p5_4578
              Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on:          March 22, 2003, 10:57:24 ; Search time 29 Seconds
                                            (without alignments)
                                            2805.459 Million cell updates/sec

Title:           US-09-784-358-2
Perfect score:   9346
Sequence:        1 MAGWIGPWWVLLLMVFMHGF..........LNGSGLGYKSRPYQS LEG 1629

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters:       112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       SwissProt_40.*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                               SUMMARIES
                 %
Result           Query
   No.   Score   Match Length DB  ID                     Description
-----------------------------------------------------------------------
     1     972    10.4   1629  1  ATS9_HUMAN             Q9p2n4 homo sapien
     2   889.5     9.5   1593  1  AT12_HUMAN             P58397 homo sapien
     3   818.5     8.8   1077  1  AT10_HUMAN             Q9h324 homo sapien
     4   593.5     6.4   1205  1  ATS2_BOVIN             P79331 b.adamts 2
     5     592     6.3   1205  1  ATS3_HUMAN             O15072 homo sapien
     6   584.5     6.3    967  1  ATS1_RAT               Q9wuq1 rattus norv
     7   571.5     6.1    967  1  ATS1_HUMAN             Q9uhi8 homo sapien
     8   564.5     6.0    968  1  ATS1_MOUSE             P97857 mus musculu
     9     563     6.0   1211  1  ATS2_HUMAN             O95450 h.adamts 2
    10   544.5     5.8    459  1  AT10_MOUSE             P58459 mus musculu
    11   529.5     5.7    860  1  ATS6_HUMAN             Q9ukp5 homo sapien
    12   496.5     5.3    890  1  ATS8_HUMAN             Q9up79 homo sapien
    13   484.5     5.2    905  1  ATS8_MOUSE             P57110 mus musculu
    14     482     5.2    930  1  ATS5_MOUSE             Q9r001 mus musculu
    15     477     5.1    930  1  ATS5_HUMAN             Q9una0 homo sapien
    16     436     4.7    997  1  ATS7_HUMAN             Q9ukp4 homo sapien
    17     415     4.4    837  1  ATS4_HUMAN             O75173 homo sapien
    18   396.5     4.2    830  1  ATS4_RAT               Q9esp7 rattus norv
    19   292.5     3.1   1522  1  BAI3_HUMAN             O60242 homo sapien
    20   289.5     3.1   1584  1  BAI1_HUMAN             O14514 homo sapien
    21     276     3.0    863  1  TSP4_XENLA             P35447 xenopus lae
    22     274     2.9    464  1  PROP_MOUSE             P11680 mus musculu
    23     271     2.9    469  1  PROP_HUMAN             P27918 homo sapien
    24     265     2.8    963  1  TSP4_RAT               P35446 rattus norv
    25   261.5     2.8   1077  1  SM5A_MOUSE             Q62217 mus musculu
    26   260.5     2.8   1093  1  SM5B_MOUSE             Q60519 mus musculu
    27   257.5     2.8   1074  1  SM5A_HUMAN             Q13591 homo sapien
    28   256.5     2.7   3707  1  PGBM_MOUSE             Q05793 mus musculu
    29     253     2.7   2012  1  DSCA_HUMAN             O60469 homo sapien
    30   251.5     2.7   1914  1  KMLS_HUMAN             Q15746 homo sapien
    31     251     2.7   1170  1  TSP1_BOVIN             Q28178 bos taurus
    32   249.5     2.7   1572  1  BAI2_HUMAN             O60241 homo sapien
    33     248     2.7    470  1  PROP_CAVPO             Q64181 cavia porce
    34     243     2.6   1173  1  TSP1_XENLA             P35448 xenopus lae
    35   242.5     2.6   1178  1  TSP2_CHICK             P35440 gallus gall
    36   241.5     2.6   1170  1  TSP2_BOVIN             Q95116 bos taurus
    37     241     2.6   5376  1  ZAN_MOUSE              O88799 mus musculu
    38     240     2.6   1170  1  TSP1_HUMAN             P07996 homo sapien
    39     240     2.6   1172  1  TSP2_HUMAN             P35442 homo sapien
    40     235     2.5   1172  1  TSP2_MOUSE             Q03350 mus musculu
    41     234     2.5   1170  1  TSP1_MOUSE             P35441 mus musculu
    42   233.5     2.5   4393  1  PGBM_HUMAN             P98160 homo sapien
    43   217.5     2.3    837  1  NCM2_MOUSE             O35136 mus musculu
    44   211.5     2.3   3133  1  HMC1_BOMMO             P98092 bombyx mori
    45   207.5     2.2   1040  1  AXO1_RAT               P22063 rattus norv
```

```
                               ALIGNMENTS


RESULT 1
ATS9_HUMAN
ID   ATS9_HUMAN     STANDARD;      PRT;  1629 AA.
AC   Q9P2N4; Q9NR29;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   ADAMTS 9 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase
DE   with thrombospondin motifs 9) (ADAM-TS 9) (ADAM-TS9).
GN   ADAMTS9 OR KIAA1312.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (SHORT ISOFORM).
RC   TISSUE=Fetal;
RX   MEDLINE=20396138; PubMed=10936055;
RA   Clark M.E., Kelner G.S., Turbeville L.A., Boyer A., Arden K.A.,
RA   Maki R.A.;
RT   "ADAMTS9, a novel member of the ADAM-TS/Metallospondin gene
RT   family.";
RL   Genomics 67:343-350(2000).
RN   [2]
RP   SEQUENCE OF 159-1629 FROM N.A. (LONG ISOFORM).
RC   TISSUE=Brain;
RX   MEDLINE=20181126; PubMed=10718198;
RA   Nagase T., Kikuno R., Ishikawa K.-I., Hirosawa M., Ohara O.;
RT   "Prediction of the coding sequences of unidentified human genes. XVI.
RT   The complete sequences of 150 new cDNA clones from brain which code
RT   for large proteins in vitro.";
RL   DNA Res. 7:65-73(2000).
CC   -!- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
CC       matrix (By similarity).
CC   -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
CC       SHORT FORM; MAY BE PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN ALL FETAL TISSUES.
CC       EXPRESSED SLIGHTLY IN ADULT OVARY, PANCREAS, HEART, KIDNEY, LUNG,
CC       PLACENTA. ALSO DETECTED IN SPINAL CORD AND BRAIN. NOT DETECTED IN
CC       COLON, SMALL INTESTINE, TESTIS, LIVER, SKELETAL MUSCLE, SPLEEN OR
CC       THYMUS.
CC   -!- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE-1 DOMAINS ARE IMPORTANT
CC       FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX (BY
CC       SIMILARITY).
CC   -!- PTM: THE PRECURSOR IS CLEAVED BY A FURIN ENDOPEPTIDASE (BY
CC       SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12B.
CC   -!- SIMILARITY: CONTAINS 1 DISINTEGRIN-LIKE DOMAIN.
CC   -!- SIMILARITY: CONTAINS 14 TSP TYPE-1 DOMAINS.
CC   --------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; AF261918; AAF89106.1; -.
DR   EMBL; AB037733; BAA92550.1; -.
DR   HSSP; P15167; 1ATL.
DR   MEROPS; M12.021; -.
DR   Genew; HGNC:13202; ADAMTS9.
DR   MIM; 605421; -.
DR   InterPro; IPR001762; Disintegrin.
DR   InterPro; IPR002870; Pep_M12B_propep.
DR   InterPro; IPR001590; Reprolysin.
DR   InterPro; IPR000884; TSP1.
DR   InterPro; IPR000130; Zn_MTpeptdse.
DR   Pfam; PF00090; tsp_1; 11.
DR   Pfam; PF01421; Reprolysin; 1.
DR   Pfam; PF01562; Pep_M12B_propep; 1.
DR   SMART; SM00209; TSP1; 12.
DR   PROSITE; PS50215; ADAM_MEPRO; 1.
DR   PROSITE; PS00427; DISINTEGRIN_1, FALSE_NEG
DR   PROSITE; PS50092; TSP1; 9.
DR   PROSITE; PS00142; ZINC_PROTEASE; 1.
KW   Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
KW   Repeat; Extracellular matrix; Alternative splicing
FT   SIGNAL        1     18       POTENTIAL.
FT   PROPEP       19    287       BY SIMILARITY.
FT   CHAIN       288   1629       ADAMTS-9.
FT   DOMAIN      509    587       DISINTEGRIN-LIKE.
FT   DOMAIN      589    642       TSP TYPE-1 1.
FT   DOMAIN      645    752       CYS RICH.
FT   DOMAIN      753    880       SPACER.
FT   DOMAIN      999   1053       TSP TYPE 1 2.
FT   DOMAIN     1056   1108       TSP TYPE-1 3.
FT   DOMAIN     1111   1156       TSP TYPE-1 4.
FT   DOMAIN     1184   1239       TSP TYPE-1 5.
FT   DOMAIN     1240   1295       TSP TYPE 1 6.
FT   DOMAIN     1332   1383       TSP TYPE-1 7.
FT   DOMAIN     1386   1439       TSP TYPE 1 8.
FT   DOMAIN     1445   1498       TSP TYPE-1 9.
FT   DOMAIN     1501   1554       TSP TYPE-1 10.
FT   DOMAIN     1562   1612       TSP TYPE-1 11.
FT   DOMAIN       88     96       POLY SER
FT   SITE        223    223       CYSTEINE SWITCH (POTENTIAL).
FT   METAL       434    434       ZINC (CATALYTIC) (BY SIMILARITY).
FT   ACT_SITE    435    435       BY SIMILARITY.
FT   METAL       438    438       ZINC (CATALYTIC) (BY SIMILARITY)
FT   METAL       444    444       ZINC (CATALYTIC) (BY SIMILARITY).
FT   CARBOHYD    112    112       N-LINKED (GLCNAC. . .) (POTENTIAL)
FT   CARBOHYD    135    135       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    271    271       N LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    749    749       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    940    940       N-LINKED (GLCNAC. . .) (POTENTIAL)
FT   CARBOHYD   1213   1213       N LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1267   1267       N LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC   1064   1072       CLVTCGKCH -> VRWEGCYFP (IN
FT                                SHORT ISOFORM).
FT   VARSPLIC   1073   1629       MISSING (IN SHORT ISOFORM).
FT   CONFLICT    367    367       F -> L (IN REF. 1).
SQ   SEQUENCE   1629 AA;  182649 MW;  C1C4CEFF58B8941F CRC64;

  Query Match             10.4%;  Score 972;  DB 1;  Length 1629;
  Best Local Similarity   19.3%;  Pred. No. 9e-53;
  Matches  326;  Conservative  145;  Mismatches  405;  Indels  814;  Gaps   56;

Qy         75 DGNWDAWGDWSDCSRTCGGGASYSLRRCLT------CRNCEGQNIRYKTCSNHDCPPDAE 128
              ||:|  :|   :   |||||||    ::| |          |: | |: :::|:|:    |
Db        588 DGSWGSWSPFGTCSRTCGGGIKTAIREGNRPEPRNGGKYCVGRRMKFKSCNIEPCLKQKR 647

Qy        129 DFRAQQCSAYNDVQYQGHYY----------EWLPRYNDPAAP--CALKCHAQGQNLVVELA 177
              ||| :||:         : ' ::            |:|:|:          | | |    |     :|
Db        648 DFRDEQCA-----HFDGKHFNINGLLPNVRWVPKYSGILMKDRCKLFCRVAGNIAYYQLR 702

Qy        178 PKVLDGTRCNTDSLDMCISGICQAVGCDRQLGSNAKEDNCGVCAGDGSTCRLVRGQSKS- 236
               :|:|||  |   |: |:|: |:|:  |||  | | |: : ||||  || |:|:  | |     :
Db        703 DRVIDGTPCGQDTNDICVQGICRQAGCDHVLNSKARRDKCGVCGGDNSSCKTVAGTFNTV 762

Qy        237 HVSPEEEEENVIAVEEGSRSVRIIVEGPAHEPESK------IIQDISRGEHSFNSPGVFV 290
              |               |: :| |: :: :           | |             |  ||||   |  |  ||
Db        763 HYG----YNTVVRIPAGATNIDVR----QHSFSGETDDDNYLALSSSKGEFLLN--GNFV 812

Qy        291 VE            NTIVEFQRGSE -      RQIFKIFGILMAEFIF          KTRYI --A 325
              |             |  ||:  |||            | :|    |:     :                 ||:
Db        813 VIMAKKEIKIGNAVVEYS GSETAVERINSTDRIEQELLLQVLSVGKLYNPDVRYSFNIP 871

Qy        326 AKDSVVQFFF-----YQPISH-----------      ---------   -  -     --      -  341
               :|    ||::         :|   |
Db        872 IEDKPQQFTWNSHGPWQAC[illegible]PGHITEPC 931

Qy        342 ------QWRQTDFFPCTVTCGGGYQLNSAEC-----VDIRLKRVVPDHYCHYYPENVKPK 390
                     :|          |:   ||  ||:        |         :|  :  ::  |  |  :|    :       ||
Db        932 GTPCHLFWHVASRSECSAQCGLLYRILTIYCAN[illegible]      LE 985

Qy        391 PKLKECSMDPCPSSDGFKEIMPYDHFQPLPRWEHNPWTACSVSCGGGIQRRSFVCVEESM 450
              |   :|        |  :                               | ::  ||  ||  ||  ||  |||    :||     :
Db        986 FSNKEKCSGEINIGG----------------   WRYSAWI[illegible]SRSCDGGTQRRRAICV   NF 1027

Qy        451 HGEILQVEEWKCMYAPKPKVMQTCNLFDCPKWIAMEWSQCTVTCGRGLRYRVVLC----- 505
                ::|   ::  || :    |    :  |  |:  |  ||:|  :  :||:|  ||||:|  ::|  |  |
Db       1028 RNDVL--DDSKCIHQEKVII-QRCSEFPCPQWKSCDWSECLVTCGKGHKHRQVWCQFGED 1084

Qy        506 -TNHRCFHVGCCNPQLKTRIKFEECVIPIPCYKPKEKSPVEAKLPWLKQAQELEETRIATE 564
                :|  |          |:|:  |          :  |    |                                                  |
Db       1085 RLNDR --   MCDPETKPTSMQTCQQP-----------------------------   -E 1107

Qy        565 EPTFIPEPWSACSTTCGPGVQVREVKCRVLLTFTQTETELPEEECEGPKLPTE-RPCLLE 623
                ::      ||     ||   |||   |   |:|   |||   ::  |:                  :   :   :|                ||:   :   |    |
Db       1108 CASWQAGPWGQCSVTCGQGYQLRAVKC  IIGTYMSV---VDDNDCNAATRPTDTQDCELP 1163

Qy        624 ACDESPASRELDIPLPEDSETTY-----DWEYAGFTPCTATCLGGHQEAIAVCLHIQTQQ 678
              :|      ||:                ||        :||                |  :       :|||:|||       |  :                    |         :
Db       1164 SCHPPPAA-------PETRRSTYSAPRTQWRFGSWTPCSATCGKGTRMRYVSC-RDENGS 1215

Qy        679 TVNDSLCDMVHRPPAMSQACNTEPCPPRWHVGSWGPCSATCGVGIQTRDVYCLHPGETPA 738
                ::|   |      :    ||    |      :    |:       ||        :|                   |          ||      |||    |           ||     |    |  ::          :
Db       1216 VADESACAILPRPVA KELCSVTPC[illegible]QWKALDWSSCSVTCGQGRATRQVMCVNYSDHVI 1273

Qy        739 PPPPC-RPEKPHALQACNQFDCPP-----------  -----    ----    ---            GWH 764
                      ||   :|      |             |    |:           ||                                                                                 |
Db       1274 [illegible]VLGGNQWR 1333

Qy        765 IEEWQQCSRTCGGGTQNRRVTCRQLLTDGSFLNLSDELCQGPKASSHKSCARTDCPPHLA 824
                   |        ||     ||    ||:|   |    |    |    |:                :|            |               :           :           !                ::|                   |      !          |
Db       1334 TGPWGACSSTCAGGSQRRVVVCQD  ENGYTAN    DCVERIKPDEQRACESGPC  PQWA 1387

Qy        825 VGEWSKCSVTCGVIIQEEKLQVQRLAAPGRRSEGESLMM[illegible]884
                 |:|   :|:       ||    ||:   |         ||||        :     |      |     |                                 ||                          ||:
Db       1388 YGNWGECTRLCGGGIRTRLVVCQR   SNGERLP           GL           SEL                  1423

Qy        885 SEMKTFIGEQGFQTT[illegible]VQEVYIQTRKHKRINLTIIGSRAYLLPNIGVIIKCPVKRFQKSLI 944

Db       1424 ------------------------------------------------------------ 1423

Qy        945 QWEKDGRCLQNSKRLGITKSGSLKIHGLAAPDIGVYRCIAGSAQETVVLKLIGTDNRLIA 1004

Db       1424 --------   ----   ----   ---------------     ---  --------  ---       -- 1423

Qy       1005 RPALREPMKEYPGMDHSEANSLGVTWHKMRQMWNNKN[illegible]SNQPPLRALLGHCS 1064

Db       1424 --     ----------------   -----      ---     --       --     -        ---     -- 1423

Qy       1065 NSAGSTNSWELKNKQFFAAVKQGAYSMDTAQFDELTRNMSQLMETGEVSDDLASQLIYQL 1124

Db       1424 ------------------------------------------------------------ 1423

Qy       1125 VAELAKAQPTHMQWRGIQEETPPAAQLRGETGSVSQSSHAKNSGKLTFKPKGPVLMRQSQ 1184
```

```
Db       1424                                                                1423

Qy       1185 PPSISFNKIINCPFNIVYIIKPIEVINIDSLLTRGDAIYIWKISIRLQSVKIIIDC 1244
                                                                    |||
Db       1424                                                       HLD  1426

Qy       1245 DGKIQIQNPIEKEQGIYPCSVANHLSSDVESSSVLYAEAPVILSVERNIIKPEHNHLSVV 1304
                        |   :| :       :
Db       1427           KPPDREQ      CN                                    1435

Qy       1305 V[illegible]IVEAALIANVIIP[illegible]                             1364
                                                                          |
Db       1436                                                            T 1436

Qy       1365 YV[illegible]AINA[illegible]KAVALDVEH[illegible]EWE[illegible]IVE[illegible]KEY[illegible]AINI[illegible]NQN[illegible] 1424
              : |                                                      |
Db       1437 HAC                                                      PH 1441

Qy       1425 EPPWEPGNWSHCSAICGHIGARIQRPQCVMANGQEVSEALCDHLQKPLAGFEPCNIRDCP 1484
                :  |   | ||  || :||   | : :    |:  :    :     | || ||  |   |    ||
Db       1442 DAAWSIGPWSSCSVSCGR[illegible]GRKQRNVYCMAK[illegible]SHLESDYCKHLAKP[illegible]RGCR[illegible]P 1499

Qy       1485 ARWFISVWSQCSVSCGDGYHSRQVICKRIKANGIVQVVSPRACAPKDRPIGRKPCPGHPC 1544
               :|     |||||||| |    | | | :         || : :      | |   ||    : | |  |
Db       1500  KWKAGAWSQCSVSCGRGVQQRHVGCQ      IGIHKIARETECNPYIRPESERDCQGPRC 1554

Qy       1545   VQW   EPGNRCPGRCMGRAVRMQQRHIACQHNSSD       SNCDDRKRPIIRRNCTSGA 1596
                  |   |     |     | |    |    |    |  : : :      : ||   |||   | : | :
Db       1555 PLYIWEAPEWL[illegible]                                           [illegible]

Qy       1597 CDVCWHTGPW 1606
              |:   | || |
Db       1612 CEYVWITGEW 1621


RESULT 2
ATL2_HUMAN
ID   ATL2_HUMAN     STANDARD;      PRT;   1593 AA.
AC   P58397;
DT   15-JUN-2002 (Rel. 41, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   ADAMTS 12 precursor (EC 3.4.24.-) (A disintegrin and
DE   metalloproteinase with thrombospondin motifs 12) (ADAM-TS 12) (ADAM-
DE   TS12).
GN   ADAMTS12.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Fetal lung;
RX   MEDLINE=21261577; PubMed=11279086;
RA   Cal S., Arguelles J.M., Fernandez P.L., Lopez-Otin C.;
RT   "Identification, characterization, and intracellular processing of
RT   ADAM-TS12, a novel human disintegrin with a complex structural
RT   organization involving multiple thrombospondin-1 repeats.";
RL   J. Biol. Chem. 276:17932-17940(2001).
CC   -!- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
CC       matrix (By similarity).
CC   -!- TISSUE SPECIFICITY: Expressed exclusively in fetal lung. Is widely
CC       expressed in gastric carcinomas and in cancer cells of diverse
CC       origin.
CC   -!- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE-1 DOMAINS ARE IMPORTANT
CC       FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX (BY
CC       SIMILARITY).
CC   -!- PTM: THE PRECURSOR IS CLEAVED BY A FURIN ENDOPEPTIDASE.
CC   -!- PTM: IS SUBJECTED TO AN INTRACELLULAR MATURATION PROCESS LEADING
CC       TO A FRAGMENT CONTAINING THE N-TERMINAL REGION INCLUDING THE
CC       METALLOPROTEINASE, DISINTEGRIN-LIKE, CYS-RICH AND TS-1 DOMAINS AND
CC       THE C-TERMINAL FRAGMENT CONTAINING THE SPACER 2 AND THE [illegible]
CC       DOMAINS.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12B.
CC   -!- SIMILARITY: CONTAINS 1 DISINTEGRIN-LIKE DOMAIN.
CC   -!- SIMILARITY: CONTAINS 8 TSP TYPE-1 DOMAINS.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AJ250725; CAC20419.1; -.
DR   Genew; HGNC:14605; ADAMTS12.
DR   MIM; 606184; -.
DR   InterPro; IPR001762; Disintegrin.
DR   InterPro; IPR002870; Pep_M12B_propep.
DR   InterPro; IPR001590; Reprolysin.
DR   InterPro; IPR000884; TSP1.
DR   InterPro; IPR000130; Zn_MIpeptdse.
DR   Pfam; PF00090; tsp_1; 6.
DR   Pfam; PF01421; Reprolysin; 1.
DR   Pfam; PF01562; Pep_M12B_propep; 1.
DR   SMART; SM00209; TSP1; 8.
DR   PROSITE; PS50215; ADAM_MEPRO; 1.
DR   PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
DR   PROSITE; PS50092; TSP1; 2.
DR   PROSITE; PS00142; ZINC_PROTEASE; 1.
KW   Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
KW   Repeat; Extracellular matrix.
FT   SIGNAL        1     25       POTENTIAL.
FT   PROPEP       26    240       BY SIMILARITY.
FT   CHAIN       241   1593       ADAMTS 12.
FT   DOMAIN      465    544       DISINTEGRIN-LIKE.
FT   DOMAIN      545    596       TSP TYPE-1 1.
FT   DOMAIN      597    700       CYS-RICH.
FT   DOMAIN      701    826       SPACER 1.
FT   DOMAIN      827    881       TSP TYPE-1 2.
FT   DOMAIN      886    943       TSP TYPE-1 3.
FT   DOMAIN      947    995       TSP TYPE-1 4.
FT   DOMAIN      996   1315       SPACER 2.
FT   DOMAIN     1316   1364       TSP TYPE-1 5.
FT   DOMAIN     1367   1423       TSP TYPE-1 6.
FT   DOMAIN     1426   1471       TSP TYPE-1 7.
FT   DOMAIN     1426   1471       TSP TYPE-1 8.
FT   DOMAIN      302    305       POLY-GLU.
FT   SITE        208    208       CYSTEINE SWITCH (POTENTIAL).
FT   METAL       392    392       ZINC (CATALYTIC) (BY SIMILARITY).
FT   ACT_SITE    393    393       BY SIMILARITY.
FT   METAL       396    396       ZINC (CATALYTIC) (BY SIMILARITY).
FT   METAL       402    402       ZINC (CATALYTIC) (BY SIMILARITY).
FT   CARBOHYD    105    105       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    125    125       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    215    215       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    485    485       N-LINKED (GLCNAC...) (POTENTIAL)
FT   CARBOHYD    685    685       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    790    790       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    951    951       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD   1104   1104       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD   1275   1275       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD   1300   1300       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD   1320   1320       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD   1371   1371       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD   1378   1378       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD   1565   1565       N-LINKED (GLCNAC...) (POTENTIAL).
SQ   SEQUENCE   1593 AA;  177545 MW;  07F9F48E63BD83A3 CRC64;

  Query Match             9.5%;  Score 889.5;  DB 1;  Length 1593;
  Best Local Similarity   24.3%;  Pred. No. 1.3e-47;
  Matches  278;  Conservative  136;  Mismatches  391;  Indels  339;  Gaps   42;

Qy         [illegible]
```

```
                  :  |    | |  |  || ||||||  |     :  | |           |: |  | :  ||: |: |
Db      536 KKPESIPGGWGRWSPWSHCSRTCGAGVQSAERLCNNPEPKFGGKYCTGERKRYRLCNVHP 595

Qy      123 CHTDAEEFRAQGTFAYNDVLYLGHTTEWLPPYNGFAAPCALKCHAQGQNLVVELAPKVLD 182
            |   :|   ||   |||  :: |  | |:       | |  |  :| || || |  |             ::    |:|
Db      596 CRSFAPTFRQMQCSFFDIVPYKNELYHWFPIFN-PAHPCELYCRPIDGQFSEKMLDAVID 654

Qy      183 GIRC--NTDSFDMCISGICQAVGCDRQLGSNAKEDNCGVCAGDGSTCRLVRGQSKSHVSP 240
            || |      :|  ::||:|||: ||||  :: ||| || |||| ||||:|: ||     |
Db      655 GIPCFEGGNSRNVCINGICKMVGCDYRIDSNATEDRCGVCLGDGSSCQIVRKMFK      709

Qy      241 EKREENVI   AVPLGSRSVRITVKGPAHLFIESKTLQGSKGEHSFNSPGVFVVE   - 292
            :|       :     :| |:| :|:          |   |:   :       |:      :     | |:::
Db      710 QKPGSGYVDIGLIPKGARDIRVMEIEGAGNFLAIR     FEDPEKYYLNGGFIIQWNGNY 765

Qy      293    NTIVPEQRGSPRQTEKIRGP----IMAEKIE-------KTRYIAAK-----SVVQFF 334
               |   :: |   :  :        ||       :        :|           |   ||   ||         | |  :
Db      766 KLAGIVPQYIPKGSIEKLMAIGPINESVWIQLLPQVINPGIKYLYIIQKDGLDNDVEQMY 825

Qy      335 FYQPISHQWRQTDFFPCTVTCGGGYQLNSAECVDIRLKRVVPDHYCHYYPENVKPKPKLK 394
            |          |:    :   |:|||| |  :     :| |:    :   :  :|      :|     ||    :|    :  |
Db      826 F--------WQYCHWTECSVTCGTGIRRQTAHCIK-KGRGMVKATFCD   PE-TQPNGRQK 874

Qy      395 ECSMDPCPSSDGFKEIMPYDHFQPLPRWEHNPWTACSVSCG-GGIQRRSFVCVEESMHGE 453
            :|      ||                          |||      |  |||  :||    |   ::|:  :|:  ::|    :
Db      875 KCHEKACP------------------PRWWAGEWEACSATCGPHGEKKRTVLCI QTMVSD 916

Qy      454 ILQVEEWKCMYAPKPKVMQTCNL-FDCPK-WIAMEWSQCTVTCGRGLRYRVVLCINHRGE 511
                 :       | :   ||| :  :||       ||   |       ||:|:|:|| |:| | | |    :   |
Db      917 EQALPPTDCQHLLKPKTLLSCNRDILCPSDWTVGNWSECSVSCGGGVRIRSVTCAKNHDE 976

Qy      512 HVGGCNPQLKLHIKEEC-------------------       --    VIPIPCYKPK      538
                 |:     | :  :   |                                       :|  |    :|:
Db      977    PCDVTRKPNSRAICGLQQCPSSRRVLKPNKGTISNGKNPPTLKPVPPPTSRPRMLTT 1033

Qy      539 -----------------EKSPVEAK-------LPWLKQAQELE------------------- 557
                               ||    |               |     :  :  |
Db     1034 PTGPESMSTSTPAISSPSPTTASKEGDLGGKQWQDSSTQPELSSRYLISTGSTSQPILTS 1093

Qy      558 --------ETRIATEEPTFIPEPWSACSTTCGPGV-QVREVKCRVLLTFTQTETELPE-- 606
                     |    :::  :        |           :|| |  | |:    |               :  |     |   | :  ||
Db     1094 QSLSIQPSEENVSSSDTGPTSEGGLVATTTSGSGLSSSRNPITWPVTPFYNILTKGPEME 1153

Qy      607 ------EECEGPK----------------------------------------------- 613
                  || |  |:
Db     1154 IHSGSGEEREQPEDKDESNPVIWTKIRVPGNDAPVESTEMPLAPPLTPDLSRESWWPPFS 1213

Qy      614 ------LPTERPCLLEA---------CDESPAS----------------------RELDIP-- 637
                  ||::||      |                | ||:                                  |  :|
Db     1214 TVMEGLLPSQRPTTSETGTPRVEGMVTEKPANTLLPLGGDHQPEPSGKTANRNHLKLPNN 1273

Qy      638 ------------LPEDSETTY----------------------------DWEYAGFTPCTATC-LGGH 664
                      |  |:   |:                                |        :: |:  ||  || :
Db     1274 MNQTKSSEPVLTEEDATSLITEGFLLNASNYKQLTNGHGSAHWIVGNWSECSTTCGLGAY 1333

Qy      665 QEAIAVCIHIQTQQTVNDSLCDMVHRP-PAMSQACNTEPCPPRWHVGSWGPCSATCGVGI 723
             :  :  |               |  ||  |   :  || ||    :  |:   ||      |  ||:|    ||    |    |
Db     1334 WKRVE-C-------ITQMDSDCAAIQRPDPA--KRCHLRPC-ACWKVGNWSKCSRNCSGGF 1383

Qy      724 QTRDVYCL----HPGETPAPPEECRDEKPHALQACNQFDCPPGWHIEEWQQCSRTCGGGT 779
             : |::  |:         |        |       :         |           :|          |        | :|   |  ||||:||||
Db     1384 KIREIQCVDSRDHRNLRPFHCQFLAGIPPPLSMSCNPEPC-EAWQVEPWSQCSRSCGGGV 1442

Qy      780 QNRRVTCRQLLTDGSFLNLSDELCQGPKASSHKSCARTDCPPHLAVGDWSKCSVSCGVGI 839
            |  |  |      |  |  :          |       |             :  |   |                  |   |  |:|    ||  ||||   |
Db     1443 QERGVECPGGLCDWTKRPTSTMSC-----NEHLCC          HWATGNWDLCSTSCGGGF 1490

Qy      840 QPRKLVCQRLAAKCEEI-PLSEMMCRHHGHRPVRRGLMRPRKPGRM-KIRLLELGR 896
            |:|     :  |   :  ::|  :                 :  :|      |     |     :   |         |  |            :     |  ||
Db     1491 QKR--IVQCVPSFGNKTEDQDQCLCDHKPRPPEFKKCNQQACKKSADLLCTKDKLSASPC 1518

Qy      897 QII-SVQRVYIQTPEEKKINLTIGSRAYLLPNTSVIIKCPVRRFQKSLIQWEKDGRCLQN 955
            |  |  ::::     :  |                                               ||           :    |  ::    |  ||
Db     1549 QTLKAMKKCSVPT-------------------VRAECCFSCPQTHIIHI--QRQRRQRLLQK 1589

Qy      956 SKRL 959
            || |
Db     1590 SKEL 1593
```

```
RESULT 3
AT10_HUMAN
ID   AT10_HUMAN      STANDARD;      PRT;  1077 AA.
AC   Q9H324;
DT   15-JUN-2002 (Rel. 41, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   ADAMTS 10 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase
DE   with thrombospondin motifs 10) (ADAM-TS 10) (ADAM-TS10) (Fragment).
GN   ADAMTS10.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Apte S.S.;
RT   "ADAM-TS10: a novel member of the ADAM-TS family containing multiple
RT   thrombospondin type 1 repeats.";
RL   Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
CC       matrix (By similarity).
CC   -!- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE 1 DOMAINS ARE IMPORTANT
CC       FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX (By
CC       similarity).
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12B.
CC   -!- SIMILARITY: CONTAINS 1 DISINTEGRIN-LIKE DOMAIN.
CC   -!- SIMILARITY: CONTAINS 5 TSP TYPE-1 DOMAINS.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AF163762; AAG35563.1; -.
DR   MEROPS; M12.235; -.
DR   Genew; HGNC:13201; ADAMTS10.
DR   InterPro; IPR001762; Disintegrin.
DR   InterPro; IPR002870; Pep_M12B_propep.
DR   InterPro; IPR001590; Reprolysin.
DR   InterPro; IPR000884; TSP1.
DR   InterPro; IPR000130; Zn_MTpeptdse.
DR   Pfam; PF00090; tsp_1; 5.
DR   Pfam; PF01421; Reprolysin; 1.
DR   Pfam; PF01562; Pep_M12B_propep; 1.
DR   SMART; SM00209; TSP1; 5.
DR   PROSITE; PS50215; ADAM_MEPRO; 1.
DR   PROSITE; PS50092; TSP1; 2.
DR   PROSITE; PS00142; ZINC_PROTEASE; 1.
DR   PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
KW   Hydrolase; Metalloprotease; Zinc; Glycoprotein; Zymogen;
KW   Repeat; Extracellular matrix.
FT   NON_TER       1      1
FT   PROPEP       <1    207       BY SIMILARITY.
FT   CHAIN       208   1077       ADAMTS-10.
FT   METAL       366    366       ZINC (CATALYTIC) (BY SIMILARITY).
FT   ACT_SITE    367    367       BY SIMILARITY.
FT   METAL       370    370       ZINC (CATALYTIC) (BY SIMILARITY).
FT   METAL       376    376       ZINC (CATALYTIC) (BY SIMILARITY).
FT   DOMAIN      434    520       DISINTEGRIN-LIKE.
FT   DOMAIN      578    679       CYS-RICH.
FT   DOMAIN      680    802       SPACER.
FT   DOMAIN      521    577       TSP TYPE 1 1.
FT   DOMAIN      749    806       TSP TYPE 1 2.
FT   DOMAIN      862    918       TSP TYPE 1 3.
FT   DOMAIN      922    976       TSP TYPE 1 4.
FT   DOMAIN      981   1031       TSP TYPE 1 5.
FT   CARBOHYD     64     64       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    196    196       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    297    297       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    714    714       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    769    769       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    866    866       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   1077 AA;  118672 MW;  3914DE1BDCBBF587 CRC64;

  Query Match             8.8%;  Score 818.5;  DB 1;  Length 1077;
  Best Local Similarity   29.7%;  Pred. No. 2e-43;
  Matches  208;  Conservative   85;  Mismatches  194;  Indels  291;  Gaps   26;

Qy         72 EDKPGNWDAWGPWSPCSRTCGGGASYSLRRC         LLGRNCSGQNIRYKICNHDCPP 125
              |   ||  |   |   : ||||||||| |  |  |  |          : |: |  | :  |:::|:  :|||
Db        518 EGVDGAWGPWIPWGDCSRTCGGGVSSSSRHCDSPRPTIGGKYCLGERRRHRSCNIDDCPP 577

Qy        126 DAEDFRAQQCSAYNDVQYQGHYYEWLPRYNDPAAPCALKCHAQGQNLVVELAPKVLDGTR 185
               :|||  |||  ::  :  :::|  :|:|                  |::    |:| |      ,   :||
Db        578 GSQDFREVQCSPFDSIPFRGKFYKWKTYRGGGVKACSLTSLAEGFNFYTERAAAVVDGTP 637

Qy        186 CNTDSLDMCISGICQAVGCDRQLGSNAKEDNCGVCAGDGSTCRLVRGQSKSHVSPEKREE 245
              |  |:::|:|| |: ||||| |||:  :|| | || |||: |   :  |    |  ||    ;
Db        638 CRPDTVDICVSGECKHVGCDRVLGSDLREDKCRVCGGDGSACETIEGVF SPASPGAGYE 696

Qy        246 NVIAVPLGSRSVRITVKGPAHLFI                     ESKTLQGSKGEHSFNSPGVF 289
              :|: :|             ||  |:||                       ||  |:|   :        .
Db        697 DVVWIP                  KGSVHIFIQDLNLSLSHLALKGDQESLLLEGLPG     PQPHRL 743

Qy        290 VVENTTVEFQRGSER QTFKIPGPLMADFIF       KTRYTAAKDSVVQFFFYQPIS  --   340
               :  ||  :  ::|  ::  |:  :   ||: |   |         :|   |        :::     ||:
Db        744 PLAGTTFQLRQGPDQVQSLEALGPINASLIVMVLARTELPA       LRYRFNAPIARDSL 798

Qy        341    HQWRQTDFFPCTVTCGGGYQLNSAECVDIRLKRVVPDHYCHYYPENVKPKPKLKECSM 398
                  :|       :  |:   |  ||  |: : ||  :          |   |||  :  :    ||  : :  |:
Db        799 PPYSWHYAPWTKCSAQCAGGSQVQAVECRNQLDSSAVAPHYCSAHSK  LPK RQRACNT 855

Qy        399 DPCPSSDGFKEIMPYDHFQPLPRWEHNPWTACSVSCGGGIQRRSFVCVEESMHGEILQVE 458
              :|||                      | |      |: || ||  |:: || ||            ::
Db        856 EPCP                      PDWVVGNWSLCSRSCDAGVRSRSVVCQRRVSAAEEKALD 898

Qy        459 EWKCMYAPKPKVMQTCNLFDCPKWIAMEWSQCTVTCGRGLRYRVVLC   INHRGEHVGG 515
               :   |     |:|  |::  |:     | |:|  |::|:||  :||  |||:|||||     :||
Db        899 DSACPQPRPPVLEACHGPTCPPEWAALDWSECTPSCGPGLRHRVVLCKSADHR        -    951

Qy        516 CNPQLKLHIKEECVIPIPCYKPKEKSPVEAKLPWLKQAQELEETRIATEEPTFIPEPWSA 575

Db        952                                                              951

Qy        576 CSTTCGPGVQVREVKCRVLLTFTQTETELPEEECEGPKLPTERPCLLEACDESPASRELD 635
                                                   ||     |                  |||:
Db        952                                ATLPPAHC                   SPAA        963

Qy        636 IPLPEDSETTYDWEYAGFTPCTATCLGGHQEATAVCLHIQTQQTVNDSLCDMVHRPPAMS 695
                                                                              :    :
Db        964                          --          --                  KPTA I   968

Qy        696 QACNTEPCPP RWHVGSWGPCSATCGVGIQTRDVYCLHPG               -----  734
                 ||     |||  ||    |  |  |||  ||||  :  |   |   |    |  |
Db        969 MRCNLRRCPPARWVAGEWGECSAQCGVGQRQRSVRCTSHTGQASHECTEALRPPTTQQCE 1028

Qy        735     ETPAP   PPEECRDEKPHALQACNQFDCPPGWHIEEWQQCSR      -----TCGG  777
                   ::|  |    ||||:|       |           ||     :  ::|  |||          |||  |
Db       1029 AKCDSPTPGDGPEECKDVNKVAY          CP    LVLKFQFCSRAYFRQMCCKTCQG 1076

RESULT 4
ATS2_BOVIN
ID   ATS2_BOVIN     STANDARD;      PRT;  1205 AA.
AC   P79331;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   ADAMTS-2 precursor (EC 3.4.24.14) (A disintegrin and
DE   metalloproteinase with thrombospondin motifs 2) (ADAM-TS 2) (ADAM-TS2)
DE   (Procollagen I/II amino propeptide processing enzyme) (Procollagen I
DE   N-proteinase) (PC I-NP) (Procollagen N-endopeptidase) (pNPI).
GN   ADAMTS2 OR NPI.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Skin;
RX   MEDLINE=97225960; PubMed=9122202;
RA   Colige A., Li S.W., Sieron A.L., Nusgens B.V., Prockop D.J.,
RA   Lapiere C.M.;
RT   "cDNA cloning and expression of bovine procollagen I N-proteinase: a
RT   new member of the superfamily of zinc-metalloproteinases with binding
RT   sites for cells and other matrix components.";
RL   Proc. Natl. Acad. Sci. U.S.A. 94:2374-2379(1997).
RN   [2]
RP   PARTIAL SEQUENCE.
RX   MEDLINE=95348096; PubMed=7622483;
RA   Colige A., Beschin A., Samyn B., Goebels Y., Van Beeumen J.,
RA   Nusgens B.V., Lapiere C.M.;
RT   "Characterization and partial amino acid sequencing of a 107-kDa
RT   procollagen I N-proteinase purified by affinity chromatography on
RT   immobilized type XIV collagen.";
RL   J. Biol. Chem. 270:16724-16730(1995).
CC   -!- FUNCTION: CLEAVES THE PROPEPTIDES OF TYPE I AND II COLLAGEN PRIOR
CC       TO FIBRIL ASSEMBLY. DOES NOT ACT ON TYPE III COLLAGEN. MAY ALSO
CC       PLAY A ROLE IN DEVELOPMENT THAT IS INDEPENDANT OF ITS ROLE IN
CC       COLLAGEN BIOSYNTHESIS.
CC   -!- CATALYTIC ACTIVITY: Cleaves the N-propeptide of collagen chain
CC       alpha 1(I) at Pro-|-Gln and of alpha 1(II) and alpha 2(I) chains
CC       at Ala-|-Gln.
CC   -!- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).
CC   -!- SUBUNIT: MAY BELONG TO A MULTIMERIC COMPLEX. BINDS SPECIFICALLY TO
CC       COLLAGEN TYPE XIV.
CC   -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
CC       matrix (By similarity).
CC   -!- TISSUE SPECIFICITY: ENZYMATIC ACTIVITY IS DETECTED AT HIGH LEVEL
CC       IN ALL TYPE I COLLAGEN-RICH TISSUES SUCH AS SKIN, BONES, TENDONS
CC       AND AORTA AND AT LOW LEVEL IN BRAIN AND THYMUS. THE MRNA LEVELS
CC       WERE DISPROPORTIONATELY HIGH IN HEART, LIVER, RETINA AND MUSCLE.
CC   -!- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE I DOMAINS ARE IMPORTANT
CC       FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX.
CC   -!- PTM: THE N-TERMINUS IS BLOCKED.
CC   -!- PTM: THE PRECURSOR IS CLEAVED BY A FURIN ENDOPEPTIDASE (BY
CC       SIMILARITY).
CC   -!- DISEASE: DEFECTS IN ADAMTS2 ARE THE CAUSE OF DERMATOSPARAXIS, A
CC       RECESSIVELY INHERITED DISORDER CHARACTERIZED BY SEVERE SKIN
CC       FRAGILITY AND BIOCHEMICALLY BY THE PRESENCE IN SKIN OF PROCOLLAGEN
CC       INCOMPLETELY PROCESSED AT THE AMINO-TERMINUS.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12B.
CC   -!- SIMILARITY: CONTAINS 1 DISINTEGRIN-LIKE DOMAIN.
CC   -!- SIMILARITY: CONTAINS 4 TSP TYPE I DOMAINS.
CC   -!- CAUTION: HAS SOMETIMES BEEN REFERRED TO ADAMTS3.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X96389; CAA65253.1; -.
DR   MEROPS; M12.301; -.
DR   InterPro; IPR001762; Disintegrin
DR   InterPro; IPR002870; Pep_M12B_propep.
DR   InterPro; IPR001590; Reprolysin.
DR   InterPro; IPR000884; TSP1.
DR   InterPro; IPR000130; Zn_MTpeptdsc.
DR   Pfam; PF00090; tsp_1; 4.
DR   Pfam; PF01421; Reprolysin; 1.
DR   Pfam; PF01562; Pep_M12B_propep; 1.
DR   SMART; SM00209; TSP1; 4.
DR   PROSITE; PS50215; ADAM_MEPRO; 1.
DR   PROSITE; PS50092; TSP1; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; FALSE_NEG.
DR   PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
KW   Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
KW   Repeat; Collagen degradation; Extracellular matrix; Heparin-binding.
FT   SIGNAL        1     28       POTENTIAL.
FT   PROPEP       29    253       BY SIMILARITY.
FT   CHAIN       254   1205       ADAMTS-2.
FT   METAL       402    402       ZINC (CATALYTIC) (POTENTIAL).
FT   ACT_SITE    403    403       BY SIMILARITY.
FT   METAL       406    406       ZINC (CATALYTIC) (BY SIMILARITY).
FT   METAL       412    412       ZINC (CATALYTIC) (BY SIMILARITY).
FT   DOMAIN      474    554       DISINTEGRIN LIKE.
FT   DOMAIN      555    611       TSP TYPE-1 1.
FT   DOMAIN      612    716       CYS-RICH.
FT   DOMAIN      717    845       SPACER
FT   DOMAIN      846    905       TSP TYPE-1 2.
FT   DOMAIN      906    968       TSP TYPE-1 3.
FT   DOMAIN      969   1024       TSP TYPE 1 4.
FT   SITE        685    687       CELL ATTACHMENT SITE (POTENTIAL).
FT   DOMAIN       31     35       POLY-ALA.
FT   DOMAIN      177    180       POLY-GLU.
FT   CARBOHYD    104    104       N LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    215    215       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    942    942       N LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    943    943       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    987    987       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1025   1025       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1092   1092       N-LINKED (GLCNAC. . .) (POTENTIAL)
FT   CARBOHYD   1139   1139       N LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1144   1144       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   1205 AA;  133987 MW;  7B5B232A45320371 CRC64;

  Query Match             6.4%;  Score 593.5;  DB 1;  Length 1205;
  Best Local Similarity   27.9%;  Pred. No. 2.8e-29;
  Matches  178;  Conservative   86;  Mismatches  266;  Indels  109;  Gaps   27;

Qy         74 KIGNWDAWGDWSDCSRTCGGGASYSLRRCLT--      GRNCDGQNIKYKTCSNHDCPPLA 127
               :||||  ||   :   ||||||  |    :    |:|           ||  |  |       :: |:: |||
Db        554 RDGNWGAWSPFGSCSRTCGTGVKFRTRQCDNPHPANGGRTCSGLAYDFQLCNSQDCPDAL 613

Qy        128 EDFRAQQCSAYN         DVQYQGHYYEWLP-RYNDPAAPCALKCHAQGQNLVVELAPKV 180
                |||  :||   ::            |  |         :  |||    :  |        |  |  |  ::       ||  :     |
Db        614 ADFRFEQCPQWDLYFFHCDAQ-----HHWLPHEHPDAKERCHLYCESKETGEVVSMKRMV 669

Qy        181 IDGTRCN-TDSLDMCISGICQAVGCDRQLGSNAKEDNCGVCAGDGSTCRLVRGQSKSHVS 239
                |||||:    |:    ::|:  |  |:  ||||    :||:  :||  ||||  ||  |  |::|:|         |
Db        669 HDGTRCSYKDAFSLCVRGDCRKVGCDGVIGSSKQEDKCGVCGGDNSHCKVVKGTFSR--S 726

Qy        240 PEKRRE NVIAVPLGSRSVRITVKGPAHLFIESKTIQGSK    GEHSFNSPG   VFVVEN 293
               |:|        :   :|    :: :  :  |                  :   |  |:   |       |  :      |        |:
Db        727 PKKLGYIKMFEIPAGARHLLIQEADTTSHHLAVKNLETGKFILNEENDVDPNSKTFIAMG 786

Qy        294 IIVEFQRGSERQIEKIPGPLM-------------ADFIFKTRYTAAKDSV----VQFFFYQ 337
                    |::        |:|   :     :||                  |          :|       :||:
Db        787 VEWEYRDEDGRETLQTMGPLHCTITVLVIPEGDARISLTYKYMIHEDSLNVDDNNVLEDD 846

Qy        338 PISHQWRQTDFFPCTVTCGGGYQLNSAECVDIRLKRVVPDHYCHYYPENV KPKPKLKEC 396
               :  ::|          :  |||:   :||||  |                 |               ::|           :|          ::|  |||        :  |
Db        847 SVGYEWALKKWSPCSKPCGGGSQFTKYGCRRRLDHKMVHRGFC----DSVSKPKAIRRIC 902

Qy        397 SMDPCPSSIGFKEIMPYDHFQPLPRWEHNPWTACSVSCGGGIQRRSFVCVEESMHGEII 455
               :    |                                  |  |       |    ||  |||    |:|  ||    ||:        :
Db        903 NPQPCSQ                   PVWVTGEWEPCSRSCGRTGMQVRSVRCVQPLHNNTTR 946

Qy        456 QVEEWKCMYAPKPKVMQTCNLFDCP KWIAMEWSQCTVTCGRGLRYRVVLC NHRGEHVG 514
                 |         |     |  :|:    :  ||       ||   :|  |      ||||:||||  |  :  |  |||          :    |
Db        947 SVHTKHCNDA RPEGRRACNRELCPGRWRAGSWSQCSVTCGNGTQERPVLC RTADDSFG 1004

Qy        515 GCNPQLEKPKLEEVIIRIPCYE      IKEKSIVEAELPWLEQA    QLEEEI      KIAT 563
                 |   :  :            |  :   ||   :             |  :||  |      :  ||   :                 :: ||          |
Db       1005 VCREE-RPETARTCRLG-PCPRNISDPSKKSYV---VQWLSRPDPNSPVQKISSKGRCQG 1059

Qy        564 EEPTF-------------IPEPWSACSTTCGPGVQVREVKCRV --           LLTFTQTE 601
                 ::  |                             ||          |    :|  |      :  :|    |                                  |    |  :
Db       1060 DKSVFCRMEVLSRYCSIPGYNKLCCKSCNLHNNLTNVEDRAEPPSLGKHNDIEEPMPVSV 1119

Qy        602 TELPEEECEGPKLPTERP---CLLEACDESPASRELDIP 637
                  |     |          |   :|   |   |                 ::   |   :      :|:|
Db       1120 PTLVMEVQPPPGIPLEVPLNTSSTNATEDHPETNAVDVP 1158
```

```
RESULT 5
ATS3_HUMAN
ID   ATS3_HUMAN      STANDARD;      PRT;  1205 AA.
AC   O15072; Q9BXZ8;
DT   16-OCT-2001 (Rel. 40, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   ADAMTS-3 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase
DE   with thrombospondin motifs 3) (ADAM-TS 3) (ADAM-TS3) (Procollagen II
DE   amino propeptide processing enzyme) (Procollagen II N-proteinase) (PC
DE   II-NP).
GN   ADAMTS3 OR KIAA0366.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE OF 1-227 FROM N.A.
RX   MEDLINE=21402912; PubMed=11408482;
RA   Fernandes R.J., Hirohata S., Engle J.M., Colige A., Cohn D.H.,
RA   Eyre D.R., Apte S.S.;
RT   "Procollagen II amino propeptide processing by ADAMTS-3. Insights on
RT   dermatosparaxis.";
RL   J. Biol. Chem. 276:31502-31509(2001).
RN   [2]
RP   SEQUENCE OF 5-1205 FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=97349984; PubMed=9205841;
RA   Nagase T., Ishikawa K.-I., Nakajima D., Ohira M., Seki N.,
RA   Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT   "Prediction of the coding sequences of unidentified human genes. VII.
RT   The complete sequences of 100 new cDNA clones from brain which can
RT   code for large proteins in vitro.";
RL   DNA Res. 4:141-150(1997).
CC   -!- FUNCTION: Cleaves the propeptides of type II collagen prior to
CC       fibril assembly. Does not act on types I and III collagens.
CC   -!- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
CC       matrix (By similarity).
CC   -!- TISSUE SPECIFICITY: Found in cartilage and skin.
CC   -!- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE-1 DOMAINS ARE IMPORTANT
CC       FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX.
CC   -!- PTM: THE PRECURSOR IS CLEAVED BY A FURIN ENDOPEPTIDASE (BY
CC       SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12B.
CC   -!- SIMILARITY: CONTAINS 1 DISINTEGRIN LIKE DOMAIN.
CC   -!- SIMILARITY: CONTAINS 4 TSP TYPE 1 DOMAINS.
CC   -!- CAUTION: HAS SOMETIMES BEEN REFERRED TO ADAMTS4.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AF247668; AAK28400.1; -.
DR   EMBL; AB002364; BAA20821.1; -.
DR   MEROPS; M12.220; -.
DR   Genew; HGNC:219; ADAMTS3.
DR   MIM; 605011; -.
DR   InterPro; IPR001762; Disintegrin.
DR   InterPro; IPR002870; Pep_M12B_propep.
DR   InterPro; IPR001590; Reprolysin.
DR   InterPro; IPR000884; TSP1.
DR   InterPro; IPR000130; Zn_MIpeptdse.
DR   Pfam; PF00090; tsp_1; 4.
DR   Pfam; PF01421; Reprolysin; 1.
DR   Pfam; PF01562; Pep_M12B_propep; 1.
DR   SMART; SM00209; TSP1; 4.
DR   PROSITE; PS50215; ADAM_MEPRO; 1.
DR   PROSITE; PS50092; TSP1; 2.
DR   PROSITE; PS00142; ZINC_PROTEASE; FALSE_NEG.
DR   PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
KW   Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
KW   Repeat; Extracellular matrix; Heparin binding.
FT   SIGNAL        1     20       POTENTIAL.
FT   PROPEP       21    249       BY SIMILARITY.
FT   CHAIN       250   1205       ADAMTS 3.
FT   METAL       398    398       ZINC (CATALYTIC) (BY SIMILARITY).
FT   ACT_SITE    399    399       BY SIMILARITY.
FT   METAL       402    402       ZINC (CATALYTIC) (BY SIMILARITY).
FT   METAL       408    408       ZINC (CATALYTIC) (BY SIMILARITY).
FT   DOMAIN      470    550       DISINTEGRIN-LIKE.
FT   DOMAIN      551    607       TSP TYPE-1 1.
FT   DOMAIN      608    712       CYS-RICH.
FT   DOMAIN      713    844       SPACER.
FT   DOMAIN      845    902       TSP TYPE-1 2.
FT   DOMAIN      903    965       TSP TYPE-1 3.
FT   DOMAIN      966   1017       TSP TYPE-1 4.
FT   DOMAIN      246    249       POLY-ARG.
FT   CARBOHYD     83     83       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    115    115       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    242    242       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    345    345       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    475    475       N-LINKED (GLCNAC. . .) (POTENTIAL)
FT   CARBOHYD    814    814       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    942    942       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   1205 AA;  135590 MW;  EB07B286FC85FB87 CRC64;

  Query Match             5.3%;  Score 592;  DB 1;  Length 1205;
  Best Local Similarity   28.0%;  Pred. No. 3.4e-29;
  Matches  156;  Conservative   73;  Mismatches  233;  Indels   76;  Gaps   21;

Qy         36 FEEALSPQGSFLEDT     TGFQFLIYRYDDQISPNIRSDEDKDGNWDAWGDWSDCSRT-C 92
                |  |   : :|    :  :       : :        : : :||||  :    :   |: :|:|
Db        513 PYE-CFKKTTEKSCETACE      WCTKCHTMWKNANLKLIGNWTDWKRTDTQCS-CTFRS 568

Qy         93 GGASYSLRRCLT         GRNCEGQNIRYKTCSNHDCPPDAEDFRAQQCSAYND-VQYQG 145
                  |     : |:|             |::|  |  |    |: |:   :|        ||||:||     |        :|:
Db        569 TGVRFFTRQCNNPMPINGGQDCPGVNFEYQLCNTEDCQKHFEDFRAQQCQQRNSHFEYQN 628

Qy        146 HYYEWLPLKYNCTAAFTALKCHALSLNEVVELAFKVLDCTLTNGLLGMCLDGLCAIICQ 203
                  : |||   : ||      |  |  |  |  ::         |   :      |  ||||  |:    |       :|:  |   |    ||
Db        629 TFSHWLPYEHPDKKECHLYC-QSKETGDVAYMKQLVHDGTRCSYKDPYSICVRGECVKVG 688

Qy        204 CDRQLGSNAKEDNCGVCAGDGSTCKLVRGQSKSHVSPEKREE-NVIAVPLGSRSVRI--TL 260
                  |  |:::|||    ||    |||||  ||:    |    |   ||   |:||        :|   |       :   :|  |   |:    :   |
Db        689 CDKFIGSNKVEDKCGVCGGDNSHCKIVKGTFTR--TPKKLGYLKMFDIPPGARHVLIQED 746

Qy        261 VKGPAHLFIESKTLQGSKGEHSFNSPG        VFVVENTTVEFQRGSERQTFKIPGPLM 314
                        |  |:::        :   :     |       |                :   :               :   :      :   :   :    |:|
Db        747 EASPHILAIKNQ      ATGHYILNGKGEEAKSRTFIDLGVEWDYNIEDDIESLHTDGPLH 802

Qy        315 ADFI                    EKIEYIAAKDSVVQFFFYQPI      SHQWRQTLFFPCIVE  356
                    |                             :|     :  ||                   |              :   :|            :           :     |:
Db        803 EPVIVLIIPQENETRSSLTYKYIIHEDSQPTINSNNVIQEETTYEWALKSWSQVSKPCGGGSQYTKYGCRRKSDNKMVHRSFCE   862

Qy        357 GGYQLNSAECVPTRLKFVVPGDYGHYYFSNVKRKPNLKLSMDPGTSSLQGEKEIMPYDHF 416
                  ||:|           |                 ::|        :|                |  |||||      :  |::      |
Db        863 GGFQYIKYGCRRKSDNKMVHRSFCE     ANKKPKPIRRMCNIQECTH                 906

Qy        417 QPLPRWEHNPWTACSVSCGGGIQRRSFVCVEESMHGEILQVEEWKCMYAPKPKVMQTCN 475
                        |  |             |    |: :|:     |   |   |:    |::         :   |                +             ||              :|:    :   ||
Db        907 ---PLWVAEEWEHCTKTCGSSGYQLRTVRCLQPLLDGTNRSVHSKYCMGDRPESRRPCN 962

Qy        476 LFDCP-KWIAMEWSQCTVTCGRGLRYRVVLCINHRGEHVGGCNPQLKLHIKEECVIPIPC 534
                          ||  :|         ||::|:||||  |           |  ::||          |:|      :        |:                          :   :|   ||
Db        963 RVPCPAQWKTGPWSECSVTCGEGTEVRQVLC---RAGDHCDGEKPE        SVRACQLPPC 1023


RESULT 6
AIS1_RAT
ID   AIS1_RAT       STANDARD;       PRT;   967 AA.
AC   Q9WUQ1; Q9ERI1;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   ADAMTS-1 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase
DE   with thrombospondin motifs 1) (ADAM-TS 1) (ADAM-TS1).
GN   ADAMTS1.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Sprague-Dawley; TISSUE=Brain;
RA   Liu X., Tu Y., Yin T., Johnstone E.M., Stephenson D.T., Clemens J.A.,
RA   Little S.P.;
RT   "Induction of a disintegrin and metalloprotease with the
RT   thrombospondin type I motif (ADAMTS).";
RL   Submitted (MAY-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE OF 18-967 FROM N.A.
RC   STRAIN=Sprague-Dawley; TISSUE=Liver;
RX   MEDLINE=20304099; PubMed=10847486;
RA   Diamantis I., Luethi M., Boesli M., Reichen J.;
RT   "Cloning of the rat ADAMTS-1 gene and its down regulation in
RT   endothelial cells in cirrhotic rats.";
RL   Liver 20:165-172(2000).
CC   -!- FUNCTION: CLEAVES AGGRECAN, A CARTILAGE PROTEOGLYCAN, AND MAY BE
CC       INVOLVED IN ITS TURNOVER. HAS ANGIOGENIC INHIBITOR ACTIVITY (BY
CC       SIMILARITY). ACTIVE METALLOPROTEASE, WHICH MAY BE ASSOCIATED WITH
CC       VARIOUS INFLAMMATORY PROCESSES AS WELL AS DEVELOPMENT OF CANCER
CC       CACHEXIA. MAY PLAY A CRITICAL ROLE IN FOLLICULAR RUPTURE (BY
CC       SIMILARITY).
CC   -!- CATALYTIC ACTIVITY: CLEAVES AGGRECAN AT THE 1938-GLU-|-LEU-1684
CC       SITE, WITHIN THE CHONDROITIN SULFATE ATTACHMENT DOMAIN.
CC   -!- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
CC       matrix (By similarity).
CC   -!- INDUCTION: DOWN REGULATED IN ENDOTHELIAL CELLS DERIVED FROM
CC       CIRRHOTIC LIVER.
CC   -!- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE-1 DOMAINS ARE IMPORTANT
CC       FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX.
CC   -!- PTM: THE PRECURSOR IS CLEAVED BY A FURIN ENDOPEPTIDASE (BY
CC       SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12B.
CC   -!- SIMILARITY: CONTAINS 1 DISINTEGRIN-LIKE DOMAIN.
CC   -!- SIMILARITY: CONTAINS 3 TSP TYPE-1 DOMAINS.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; AF149118; AAD34012.1; -.
DR   EMBL; AF304446; AAG29823.1; -.
DR   MEROPS; M12.222; -.
DR   InterPro; IPR001762; Disintegrin.
DR   InterPro; IPR002870; Pep_M12B_propep.
DR   InterPro; IPR001590; Reprolysin.
DR   InterPro; IPR000884; TSP1.
DR   InterPro; IPR000130; Zn_MIpeptdse.
DR   Pfam; PF00090; tsp_1; 6.
DR   Pfam; PF01421; Reprolysin; 2.
DR   Pfam; PF01562; Pep_M12B_propep; 2.
DR   SMART; SM00209; TSP1; 3.
DR   PROSITE; PS50215; ADAM_MEPRO; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; 1.
DR   PROSITE; PS50092; TSP1; 2.
DR   PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
KW   Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
KW   Repeat; Extracellular matrix; Heparin-binding.
FT   SIGNAL        1     54       POTENTIAL.
FT   PROPEP       55    252       BY SIMILARITY.
FT   CHAIN       253    967       ADAMTS-1.
FT   SITE        205    205       CYSTEINE SWITCH (POTENTIAL).
FT   METAL       401    401       ZINC (CATALYTIC) (BY SIMILARITY).
FT   ACT_SITE    402    402       BY SIMILARITY.
FT   METAL       405    405       ZINC (CATALYTIC) (BY SIMILARITY).
FT   METAL       411    411       ZINC (CATALYTIC) (BY SIMILARITY).
FT   DOMAIN      476    558       DISINTEGRIN-LIKE.
FT   DOMAIN      559    615       TSP TYPE-1 1.
FT   DOMAIN      616    724       CYS-RICH.
FT   DOMAIN      725    857       SPACER.
FT   DOMAIN      858    907       TSP TYPE-1 2.
FT   DOMAIN      908    967       TSP TYPE-1 3.
FT   DOMAIN      194    198       POLY-ARG.
FT   CARBOHYD    547    547       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    720    720       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    764    764       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    782    782       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    945    945       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT     21     21       I -> V (IN REF. 2).
FT   CONFLICT     26     31       KFPSSQ -> RSRGSL (IN REF. 2).
FT   CONFLICT     49     49       V -> A (IN REF. 2).
FT   CONFLICT     72     72       R -> P (IN REF. 2).
FT   CONFLICT     79     79       L -> TR (IN REF. 2).
FT   CONFLICT    249    249       E -> G (IN REF. 2).
FT   CONFLICT    262    265       TMLV -> NLLK (IN REF. 2).
FT   CONFLICT    607    607       S -> F (IN REF. 2).
FT   CONFLICT    936    936       L -> V (IN REF. 2).
FT   CONFLICT    962    962       I -> T (IN REF. 2).
SQ   SEQUENCE   967 AA;  105705 MW;  F93C864F6DCDB4CF CRC64;

  Query Match              6.3%;  Score 584.5;  DB 1;  Length 967;
  Best Local Similarity   27.7%;  Pred. No. 7.5e-29;
  Matches  132;  Conservative   75;  Mismatches  176;  Indels   53;  Gaps   16;

Qy         76 GNWDAWGDWSDCSRTCGGGASYSLRRCLT------GRNCEGQNIRYKTCSNHDCPP-DAE 128
               |:|  || | ||||||||  |::| |          |: |||: :||::|: |||  : :
Db        560 GSWGPWGPWGDCSRTCGGGVQYIMRECDNPVPKNGGKYCEGKRVRYRSCNIEDCPDNNGK 619

Qy        129 DFFAQQCSAYNDVQYQGH----YYEWLPRYN--DPAAPCALKCHAQGQNLVVELAPKVLD 182
                || :|| |:|:                  || |:|     |    | | | |:|         | |||:|
Db        620 TFREEQCEAHNEFSKASFGNEPTVEWTPKYAGVSPKDRCKLTCEAKGIGYFFVLQPKVVD 679

Qy        183 GTRCNTDSLDMCISGICQAVGCDRQLGSNAKEDNCGVCAGDGSTCRLVRGQSKSHVSPEK 242
              || |: ||   :|: | |    |||| : |   | | |||| |:||||: : |    : |
Db        680 GTPCSPDSTSVCVQGQCVKAGCDRIIDSKKKFDKCGVCGGNGSTCKKISG---IVTSTRP 736

Qy        243 REENVIAVPLGSRSVRITVKGPAHLFIESKTLQGSKGEHSF----NSPGVFVVENTTVEF 298
                  ::: :| |: :: : : |               :||:    ||      : | :::      |
Db        737 CYHDIVTIPAGATNIEVKHRNP----------KGSRNNGSFLAIRAADGTYILNG---NF 783

Qy        299 QRGSDRQIFKIPGPLMADFIFKIRYIAAKDSVVQFFFYQPISHQWRQIDFFPCTV-TCGG 357
                  :  |       | ::          ||: :  :: :    : |:                | |:
```

```
Db        784 TLSTLEQDLTYKGTVL-------RYSGSSAALERIRSFSPLKE--------PLTIQVLMV 828

Qy        358 GYQLNSAFCVDIRLKRVVPKGYCHYYPENVKPKPKLKECSMEPCPSSDGFKEIMPYDHFQ 417
               |: |                          :||:|                |:   |::
Db        829 GHAL----------------------------RPKIKYTYFMK-------KKTEPFNAIP 853

Qy        418 PLPRWEHNPWTACSVSCGGGIQRRSFVCVEESMHGEILQVEEWKCMYAPKPKVMQTCNLF 477
                  |      |    || :|| | |||    | : : |              :|     ||    : |
Db        854 TFSEWVIEEWGECSKTCGSGWQRRVVECRDINGH------PASECAKEVKPASTRPCADL 907

Qy        478 DCPRWIAMEWSQCTVTCGRGLRYRVVLCINHRGEHVG--GCNPQLK-LHIKEECVI 530
                ||:|    :|| |: |||:| : | : |::| |  :      |:|  |  |  : |::
Db        908 PCPRWQVGDWSPCSKTCGKGYKKRTLKCLSHDGGVLSNESCDPLKKPKHYIDFCTL 963

RESULT 7
ATS1_HUMAN
ID   ATS1_HUMAN     STANDARD;      PRT;   967 AA.
AC   Q9UHI8; Q9UP80; Q9UH83; Q9P2K0; Q9NSJ8;
DT   30-MAY-2000 (Rel. 39, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   ADAMTS-1 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase
DE   with thrombospondin motifs 1) (ADAM-TS 1) (ADAM-TS1) (METH-1).
GN   ADAMTS1 OR METH1 OR KIAA1346.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Casas C., Pritchard M.A., Estivill X., Arbones M.L.;
RT   "Cloning, characterization and mapping on human chromosome 21 of the
RT   orthologue of murine Adamts-1.";
RL   Submitted (JUL 1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A., AND FUNCTION.
RC   TISSUE=Heart;
RX   MEDLINE=99367466; PubMed=10438512;
RA   Vazquez F., Hastings G., Ortega M. A., Lane T.F., Oikemus S.,
RA   Lombardo M., Iruela Arispe M.L.;
RT   "METH 1, a human ortholog of ADAMTS-1, and METH-2 are members of a new
RT   family of proteins with angio-inhibitory activity.";
RL   J. Biol. Chem. 274:23349-23357(1999).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Endothelial cells;
RX   MEDLINE=20247184; PubMed=10785405;
RA   Glienke J., Schmitt A.O., Pilarsky C., Hinzmann B., Weiss B.,
RA   Rosenthal A., Thierauch K.H.;
RT   "Differential gene expression by endothelial cells in distinct
RT   angiogenic states.";
RL   Eur. J. Biochem. 267:2820-2830(2000).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=20181126; PubMed=10718198;
RA   Nagase T., Kikuno R., Ishikawa K. I., Hirosawa M., Ohara O.;
RT   "Prediction of the coding sequences of unidentified human genes.XVI.
RT   The complete sequences of 150 new cDNA clones from brain which code
RT   for large proteins in vitro.";
RL   DNA Res. 7:65-73(2000).
RN   [5]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20289799; PubMed=10830953;
RA   Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,
RA   Park H.-S., Toyoda A., Ishii K., Totoki Y., Choi D.-K., Soeda E.,
RA   Ohki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K., Polley A.,
RA   Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,
RA   Reichwald K., Rump A., Schillhabel M., Schudy A., Zimmermann W.,
RA   Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S.,
RA   Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E.,
RA   Minoshima S., Shimizu N., Nordsiek G., Hornischer K., Brandt P.,
RA   Scharfe M., Schoen O., Desario A., Reichelt J., Kauer G., Bloecker H.,
RA   Ramser J., Beck A., Klages S., Hennig S., Riesselmann L., Dagand E.,
RA   Wehrmeyer S., Borzym K., Gardiner K., Nizetic D., Francis F.,
RA   Lehrach H., Reinhardt R., Yaspo M. L.;
RT   "The DNA sequence of human chromosome 21.";
RL   Nature 405:311-319(2000).
RN   [6]
RP   SEQUENCE OF 418-967 FROM N.A.
RC   TISSUE=Melanoma;
RA   Blum H., Bauersachs S., Mewes H. W., Weil B., Wiemann S.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: CLEAVES AGGRECAN, A CARTILAGE PROTEOGLYCAN, AND MAY BE
CC       INVOLVED IN ITS TURNOVER (BY SIMILARITY). HAS ANGIOGENIC INHIBITOR
CC       ACTIVITY. ACTIVE METALLOPROTEASE, WHICH MAY BE ASSOCIATED WITH
CC       VARIOUS INFLAMMATORY PROCESSES AS WELL AS DEVELOPMENT OF CANCER
CC       CACHEXIA. MAY PLAY A CRITICAL ROLE IN FOLLICULAR RUPTURE.
CC   -!- CATALYTIC ACTIVITY: CLEAVES AGGRECAN AT THE 1938-GLU-|-LEU-1939
CC       SITE, WITHIN THE CHONDROITIN SULFATE ATTACHMENT DOMAIN.
CC   -!- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
CC       matrix (By similarity).
CC   -!- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE-1 DOMAINS ARE IMPORTANT
CC       FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX.
CC   -!- PTM: THE PRECURSOR IS CLEAVED BY A FURIN ENDOPEPTIDASE (BY
CC       SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12B.
CC   -!- SIMILARITY: CONTAINS 1 DISINTEGRIN-LIKE DOMAIN.
CC   -!- SIMILARITY: CONTAINS 3 TSP TYPE-1 DOMAINS
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AF170084; AAF15317.1; -.
DR   EMBL; AF060152; AAD48080.1; ALT_INIT.
DR   EMBL; AF207664; AAF23772.1; -.
DR   EMBL; AB037767; BAA92584.1; ALT_INIT.
DR   EMBL; AP001697; BAA95502.1; -.
DR   EMBL; AL162080; CAB82413.1; -.
DR   MEROPS; M12.222;
DR   Genew; HGNC:217; ADAMTS1.
DR   MIM; 605174; -.
DR   InterPro; IPR001762; Disintegrin.
DR   InterPro; IPR002870; Pep_M12B_propep.
DR   InterPro; IPR001590; Reprolysin.
DR   InterPro; IPR000884; TSP1.
DR   InterPro; IPR000130; Zn_MIpeptdse.
DR   Pfam; PF00090; tsp_1; 3.
DR   Pfam; PF01421; Reprolysin; 2.
DR   Pfam; PF01562; Pep_M12B_propep; 2.
DR   SMART; SM00209; TSP1; 3.
DR   PROSITE; PS50215; ADAM_MEPRO; 1.
DR   PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
DR   PROSITE; PS50092; TSP1; 2.
DR   PROSITE; PS00142; ZINC_PROTEASE; 1.
KW   Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
KW   Repeat; Extracellular matrix; Heparin-binding.
FT   SIGNAL        1     49       POTENTIAL.
FT   PROPEP       50    252       BY SIMILARITY.
FT   CHAIN       253    967       ADAMTS-1.
FT   SITE        198    198       CYSTEINE SWITCH (POTENTIAL).
FT   METAL       401    401       ZINC (CATALYTIC) (BY SIMILARITY).
FT   ACT_SITE    402    402       BY SIMILARITY.
FT   METAL       405    405       ZINC (CATALYTIC) (BY SIMILARITY).
FT   METAL       411    411       ZINC (CATALYTIC) (BY SIMILARITY).
FT   DOMAIN      476    559       DISINTEGRIN-LIKE.
FT   DOMAIN      560    616       TSP TYPE-1 1.
FT   DOMAIN      617    724       CYS-RICH.
FT   DOMAIN      725    849       SPACER.
```

```
FT   DOMAIN      850    908       TSP TYPE-1 2.
FT   DOMAIN      909    967       TSP TYPE-1 3.
FT   DOMAIN      843    846       POLY-LYS.
FT   CARBOHYD    547    547       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    720    720       N-LINKED (GLCNAC. . .) (POTENTIAL)
FT   CARBOHYD    764    764       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT    227    227       P -> A (IN REF. 4 AND 5).
FT   CONFLICT    468    468       Q -> H (IN REF. 1).
FT   CONFLICT    561    561       S -> N (IN REF. 1).
SQ   SEQUENCE   967 AA;  105383 MW;  C1893893247A1ED1 CRC64;

  Query Match              6.1%;  Score 571.5;  DB 1;  Length 967;
  Best Local Similarity   28.2%;  Pred. No. 4.9e-28;
  Matches  136;  Conservative   77;  Mismatches  164;  Indels  105;  Gaps   19;

Qy         76 GNWDAWGDWSECSRTCGGGASYSLRRCLT        GENCEGQNIRYKTCSNHDCPP DAE 128
              |:|   || |  ||||||||   |::| |            |: :|||: :||::|:   |||   : :
Db        560 GSWSMWGPWGECSRTCGGGVQYTMRECDNPVPKNGGKYCEGKRVRYRSCNLEDCPDNNGK 619

Qy        129 DFRAQQCSAYNDVQY     QGHYYEWLPRYN   DPAAPCALKCHAQGQNLVVELAPKVLD 182
                || :||  |:|:            |     ||:|:|      |     |   |   | |:|         | |||:|
Db        620 TFREEQCEAHNEFSKASFGSGPAVEWIPKYAGVSPKDRCKLICQAKGIGYFFVLQPKVVD 679

Qy        183 GTRCNTDSLDMCISGICQAVGCDRQLGSNAKEDNCGVCAGDGSTCRLVRGQSKSHVSPEK 242
              || |: ||    :|: | |     ||||  : |   | | |||| |:|||: :  |     |   | :
Db        680 GTPCSPDSTSVCVQGQCVKAGCDRIIDSKKKFDKCGVCGGNGSTCKKISG    SVTSAKP 736

Qy        243 REENVIAVPLGSRSVRITVKGPAHLFIEGKTLQGSKGEHSF      NSPGVFVVEN      T 294
                   ::| :| |: ::                    :: :   :||:     ||          : | :::            :
Db        737 GYHDIITIPTGATNIE                VKQRNQRGSRNNGSFLAIKAADGTYILNGDYTLS 786

Qy        295 TVEFQRGSERQSFKIPGPLMADFIIK      TRYTAAKDSVVQFFFYQPISHQWRQTDFFPC 351
              |:|                              | ::|      ||:  :   :: :     : |:                    |
Db        787 TLE-------------------QDIMYKGVVLRYSGSSAALERIRSFSPLKE         --PL 821

Qy        352 TVTCGGGYQLNSAECVDIRLKRVVPDHYCHYYPENVKPKPKLKECSMDPCPSSDGFKEIM 411
              |:            |          :   ::|              | ::                    | |: |  :   |:
Db        822 TIQV---LTVGNALRPKIK------YTYFV---------KKKKESFNAIPT        -    854

Qy        412 PYDHFQPLPRWEHNPWTACSVSCGGGIQRRSFVCVEESMHGEILQVEEWKCMYAPKPKVM 471
                                     |       |  || ||   |  |||    |          ::|:                :|      ||
Db        855 -------FSAWVIEEWGECSKSCELGWQRRLVECRDINGQPAS      EGAKEVKPAST 901

Qy        472 QTCNLFDCPKWIAMEWSQCTVTCGRGLRYRVVLCINHRGEHVG   GCNPQLK LHIKEEC 528
              : |       ||:|     ||| |: |||:| :  | : |::|  |    :        |:|   |    |  : |
Db        902 RPCADHPCPQWQLGEWSSCSKTCGKGYKKRSLKCLSHDGGVLSHESCDPLKKPKHFIDFC 961

Qy        529 VI 530
               :
Db        962 TM 963


RESULT 8
ATS1_MOUSE
ID   ATS1_MOUSE     STANDARD;      PRT;   968 AA.
AC   P97857; O54768;
DT   30-MAY-2000 (Rel. 39, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   ADAMTS-1 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase
DE   with thrombospondin motifs 1) (ADAM-TS 1) (ADAM-TS1).
GN   ADAMTS1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129/SvJ;
RX   MEDLINE=98110583; PubMed=9441751;
RA   Kuno K., Iizasa H., Ohno S., Matsushima K.;
RT   "The exon/intron organization and chromosomal mapping of the mouse
RT   ADAMTS-1 gene encoding an ADAM family protein with TSP motifs.";
```

```
RL   Genomics 46:466-471(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE 97150761; PubMed 8995297;
RA   Kuno K., Kanada N., Nakashima E., Fujiki F., Ichimura F.,
RA   Matsushima K.;
RT   "Molecular cloning of a gene encoding a new type of metalloproteinase-
RT   disintegrin family protein with thrombospondin motifs as an
RT   inflammation associated gene.";
RL   J. Biol. Chem. 272:556-562(1997).
RN   [3]
RP   CHARACTERIZATION, AND MUTAGENESIS OF GLU 403.
RX   MEDLINE=99303657; PubMed=10373500,
RA   Kuno K., Terashima Y., Matsushima K.;
RT   "ADAMTS-1 is an active metalloproteinase associated with the
RT   extracellular matrix.";
RL   J. Biol. Chem. 274:18821 18826(1999).
RN   [4]
RP   FUNCTION.
RX   MEDLINE-20389568; PubMed-10930576;
RA   Kuno K., Okada Y., Kawashima H., Nakamura H., Miyasaka M.,
RA   Ohno H., Matsushima K.,
RT   "ADAMTS-1 cleaves a cartilage proteoglycan, aggrecan.",
RL   FEBS Lett. 478:241-245(2000).
RN   [5]
RP   FUNCTION, AND INDUCTION.
RX   MEDLINE=20243757; PubMed 10781075;
RA   Robker R.L., Russell D.L., Espey L.L., Lydon J.P., O'Malley B.W.,
RA   Richards J.S.;
RT   "Progesterone-regulated genes in the ovulation process: ADAMTS-1 and
RT   cathepsin L proteases.";
RL   Proc. Natl. Acad. Sci. U.S.A. 97:4689-4694(2000).
CC   -!- FUNCTION: CLEAVES AGGRECAN, A CARTILAGE PROTEOGLYCAN, AND MAY BE
CC       INVOLVED IN ITS TURNOVER. HAS ANGIOGENIC INHIBITOR ACTIVITY (BY
CC       SIMILARITY). ACTIVE METALLOPROTEASE, WHICH MAY BE ASSOCIATED WITH
CC       VARIOUS INFLAMMATORY PROCESSES AS WELL AS DEVELOPMENT OF CANCER
CC       CACHEXIA. MAY PLAY A CRITICAL ROLE IN FOLLICULAR RUPTURE (BY
CC       SIMILARITY).
CC   -!- CATALYTIC ACTIVITY: CLEAVES AGGRECAN AT THE 1691-GLU-|-LEU-1692
CC       SITE, WITHIN THE CHONDROITIN SULFATE ATTACHMENT DOMAIN.
CC   -!- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: SECRETED. ASSOCIATED WITH THE EXTRACELLULAR
CC       MATRIX.
CC   -!- INDUCTION: INDUCED IN VITRO IN COLON ADENOCARCINOMA CELLS BY
CC       INTERLEUKIN-1, OR IN VIVO IN KIDNEY AND HEART BY
CC       LIPOPOLYSACCHARIDE. ALSO INDUCED BY LH STIMULATION IN GRANULOSA
CC       CELLS OF PREOVULATORY FOLLICLES.
CC   -!- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE-1 DOMAINS ARE IMPORTANT
CC       FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX.
CC   -!- PTM: THE PRECURSOR IS CLEAVED BY A FURIN ENDOPEPTIDASE.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12B.
CC   -!- SIMILARITY: CONTAINS 1 DISINTEGRIN-LIKE DOMAIN.
CC   -!- SIMILARITY: CONTAINS 3 TSP TYPE-1 DOMAINS.
CC   -!- CAUTION: REF.2 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A
CC       FRAMESHIFT IN POSITION 7.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AB001735; BAA24501.1; ALT_INIT.
DR   EMBL; D67076; BAA11088.1; ALT_FRAME.
DR   MEROPS; M12.222; -.
DR   MGD; MGI:109249; Adamts1
DR   InterPro; IPR001762; Disintegrin.
DR   InterPro; IPR002870; Pep_M12B_propep.
DR   InterPro; IPR001590; Reprolysin.
DR   InterPro; IPR000884; TSP1.
DR   InterPro; IPR000130; Zn_MTpeptdse.
DR   Pfam; PF00090; tsp_1; 3.
DR   Pfam; PF01421; Reprolysin; 1.
DR   Pfam; PF01562, Pep_M12B_propep; 1.
DR   SMART, SM00209; TSP1; 3.
DR   PROSITE, PS50215, ADAM_MEPRO; 1.
DR   PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
DR   PROSITE; PS50092; TSP1; 3.
DR   PROSITE; PS00142; ZINC_PROTEASE; 1.
KW   Hydrolase; Metalloprotease, Zinc, Signal, Glycoprotein, Zymogen;
KW   Repeat, Extracellular matrix; Heparin-binding.
FT   SIGNAL          1     48       POTENTIAL.
FT   PROPEP         49    253
FT   CHAIN         254    968       ADAMTS-1.
FT   SITE          206    206       CYSTEINE SWITCH (POTENTIAL).
FT   METAL         402    402       ZINC (CATALYTIC) (BY SIMILARITY).
FT   ACT_SITE      403    403
FT   METAL         406    406       ZINC (CATALYTIC) (BY SIMILARITY).
FT   METAL         412    412       ZINC (CATALYTIC) (BY SIMILARITY).
FT   DOMAIN        477    559       DISINTEGRIN-LIKE.
FT   DOMAIN        560    617       TSP TYPE-1 1.
FT   DOMAIN        618    725       CYS-RICH.
FT   DOMAIN        726    850       SPACER.
FT   DOMAIN        851    909       TSP TYPE-1 2.
FT   DOMAIN        910    968       TSP TYPE-1 3.
FT   DOMAIN        195    199       POLY-ARG.
FT   CARBOHYD      548    548       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD      721    721       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD      765    765       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD      783    783       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD      946    946       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   MUTAGEN       403    403       E->Q: LOSS OF ACTIVITY.
FT   CONFLICT      335    335       N -> S (IN REF. 2).
FT   CONFLICT      425    425       I -> S (IN REF. 2).
SQ   SEQUENCE   968 AA;  105841 MW;  42EBDA55499FB6C1 CRC64;

  Query Match             6.0%;  Score 564.5;  DB 1;  Length 968;
  Best Local Similarity   27.7%;  Pred. No. 1.3e-27;
  Matches  133;  Conservative   72;  Mismatches  173;  Indels  103;  Gaps   18;

Qy         76 GNWDAWGDWSDCSRTCGGGASYSLRRCLT        GRNCEGQNIRYKTCSNHDCPP DAE 128
              |:|  || | |||||||||  |::| |        |: |||: :||::|:  |||  : :
Db        561 GSWGPWGPWGDCSRTCGGGVQYTMRE[illegible]NPVPKNGGKYCEGKRVKYRSCNIEDCPDNNGK 620

Qy        129 DFRAQQCSAYNDVQYQGH----YYEWLPRYN---DPAAPCALKCHAQGQNLVVELAPKVLD 182
               || :|| |:|:             || |:|     |    | | | |:|      | |||:|
Db        621 IFREEQCEAHNEFSKASFGNEPIVEWIPKYAGV[illegible]KDRCKLTCEAKG[illegible]GYFFV[illegible]PKVVD 680

Qy        183 [illegible] 242
              || |: ||  :|: | |   |||| : |  | | |||| |:||||: : |        |
Db        681 GTPCSPDSTSVCVQGQCVKAGCDRIIDSKKKFDKCGVCGGNGSTCKKMSG---IVTSTRP 737

Qy        243 REENVIAV[illegible]SRSVR[illegible]TVK[illegible]AHHFIHS[illegible]     N[illegible]VVVV[illegible]VLP 298
                :::  :| |: ::              :: :   :||:    ||        : | :::      |
Db        738 GYHDIVTIPAGATNIE                VKHRNQRGSRNNGSFLAIRAADGTYILNG    NF 784

Qy        299 QPGSPPQTFKIPGPIMADFIPKTRYPAAKPGVVQFTPYQ[illegible]ISHQWRQTDFFPCIV-TCGG 357
                 :  |        | ::           ||: :   :: :     : |:                | |:
Db        785 ILSILEQLLTYKGIVL          [illegible]      [illegible]LIQVLMV 829

Qy        358 [illegible] 415
              |: |                                 :|||:|                |     |      :
Db        830 [illegible]                      [illegible]                 [illegible] 849

Qy        416 FQPLP---RWEHNPWTACSVSCGGGIQRRSFVCVEESMHGEILQVEEWKCMYAPKPKVMQ 472
              |  :|        |         |   || :|| | |||    | : : |              :|        ||     :
Db        850 FNAIPTFSEWVIEEWGECSKTCGSGWQRRVVQCRDINGH------PASECAKEVKPASIR 903

Qy        473 TCNLFDCPEWIAMEWSQ[illegible] 529
                |       || |      :|| |:  |||:|  :  |  :  |::| |    :         |:|    |     |     :  |
Db        904 FCADLPCPHWQVGDWS[illegible]IDFCT 963

Qy        530 I 530
              :
```

```
Db        964 L 964


RESULT 9
AT52 HUMAN
ID   AT52 HUMAN     STANDARD;      PRT;  1211 AA.
AC   O95450;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   ADAMTS-2 precursor (EC 3.4.24.14) (A disintegrin and
DE   metalloproteinase with thrombospondin motifs 2) (ADAM-TS 2) (ADAM-TS2)
DE   (Procollagen I/II amino propeptide processing enzyme) (Procollagen I
DE   N-proteinase) (PC I-NP) (Procollagen N-endopeptidase) (pNPI)
DE   (Procollagen I/II amino propeptide processing enzyme).
GN   ADAMTS2 OR PCINP OR PCPNI.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS LPNPI AND SPNPI), AND DISEASE.
RC   TISSUE=Skin;
RX   MEDLINE=99347935; PubMed=10417273;
RA   Colige A., Sieron A.L., Li S.-W., Schwarze U., Petty E.,
RA   Wertelecki W., Wilcox W., Krakow D., Cohn D.H., Reardon W.,
RA   Byers P.H., Lapiere C.M., Prockop D.J., Nusgens B.V.;
RT   "Human Ehlers-Danlos syndrome type VII C and bovine dermatosparaxis
RT   are caused by mutations in the procollagen I N-proteinase gene.";
RL   Am. J. Hum. Genet. 65:308-317(1999).
CC   -!- FUNCTION: Cleaves the propeptides of type I and II collagen prior
CC       to fibril assembly. Does not act on type III collagen. May also
CC       play a role in development that is independant of its role in
CC       collagen biosynthesis.
CC   -!- CATALYTIC ACTIVITY: Cleaves the N-propeptide of collagen chain
CC       alpha-1(I) at Pro-|-Gln and of alpha-1(II) and alpha-2(I) chains
CC       at Ala-|-Gln.
CC   -!- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).
CC   -!- SUBUNIT: MAY BELONG TO A MULTIMERIC COMPLEX. BINDS SPECIFICALLY TO
CC       COLLAGEN TYPE XIV (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
CC       matrix (By similarity).
CC   -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; LPNPI (SHOWN HERE), AND SPNPI;
CC       ARE PRODUCED BY ALTERNATIVE SPLICING. SPNPI HAS NO SIGNIFICANT N
CC       PROCOLLAGEN PEPTIDASE ACTIVITY.
CC   -!- TISSUE SPECIFICITY: EXPRESSED AT HIGH LEVEL IN SKIN, BONE, TENDON
CC       AND AORTA AND AT LOW LEVELS IN THYMUS AND BRAIN.
CC   -!- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE-1 DOMAINS ARE IMPORTANT
CC       FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX.
CC   -!- PTM: THE PRECURSOR IS CLEAVED BY A FURIN ENDOPEPTIDASE (BY
CC       SIMILARITY).
CC   -!- DISEASE: Defects in ADAMTS2 are the cause of Ehlers-Danlos
CC       syndrome type VIIC (EDS VIIC), a recessively inherited connective-
CC       tissue disorder characterized clinically by severe skin fragility
CC       and joint hypermobility and biochemically by the presence in skin
CC       of procollagen incompletely processed at the N-terminus.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12B.
CC   -!- SIMILARITY: CONTAINS 1 DISINTEGRIN-LIKE DOMAIN.
CC   -!- SIMILARITY: CONTAINS 4 TSP TYPE-1 DOMAINS.
CC   -!- CAUTION: HAS SOMETIMES BEEN REFERRED TO ADAMTS3.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AJ003125; CAA05880.1; -.
DR   MEROPS; M12.301; -.
DR   Genew; HGNC:218; ADAMTS2.
DR   MIM; 604539; -.
DR   MIM; 225410; -.
DR   InterPro; IPR001762; Disintegrin.
DR   InterPro; IPR002870; Pep_M12B_propep.
DR   InterPro; IPR001590; Reprolysin.
DR   InterPro; IPR000884; TSP1.
DR   InterPro; IPR000130; Zn_MIpeptdse.
DR   Pfam; PF00090; tsp_1; 4.
DR   Pfam; PF01421; Reprolysin; 1.
DR   Pfam; PF01562; Pep_M12B_propep; 1.
DR   SMART; SM00209; TSP1; 4.
DR   PROSITE; PS50215; ADAM_MEPRO; 1.
DR   PROSITE; PS50092; TSP1; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; FALSE_NEG.
DR   PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
KW   Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
KW   Repeat; Collagen degradation; Extracellular matrix; Heparin-binding;
KW   Alternative splicing; Ehlers-Danlos syndrome.
FT   SIGNAL        1     29       POTENTIAL.
FT   PROPEP       30    253       BY SIMILARITY.
FT   CHAIN       254   1211       ADAMTS 2.
FT   METAL       408    408       ZINC (CATALYTIC) (BY SIMILARITY).
FT   ACT_SITE    409    409       BY SIMILARITY.
FT   METAL       412    412       ZINC (CATALYTIC) (BY SIMILARITY).
FT   METAL       418    418       ZINC (CATALYTIC) (BY SIMILARITY).
FT   DOMAIN      480    560       DISINTEGRIN-LIKE.
FT   DOMAIN      561    617       TSP TYPE-1 1.
FT   DOMAIN      618    722       CYS-RICH.
FT   SITE        691    693       CELL ATTACHMENT SITE (POTENTIAL).
FT   DOMAIN      723    851       SPACER.
FT   DOMAIN      852    911       TSP TYPE-1 2.
FT   DOMAIN      912    974       TSP TYPE-1 3.
FT   DOMAIN      975   1030       TSP TYPE-1 4.
FT   DOMAIN       40     43       POLY-ALA.
FT   DOMAIN      185    188       POLY-GLU.
FT   CARBOHYD    112    112       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    251    251       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    949    949       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    993    993       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1031   1031       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1098   1098       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1145   1145       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1150   1150       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC    544    566       HCRKNHLWGAFRLRKRRSSWGA -> FRKGAVARAQYES
FT                                TLSGQGRWTA (IN ISOFORM SPNPI).
FT   VARSPLIC    567   1211       MISSING (IN ISOFORM SPNPI).
SQ   SEQUENCE   1211 AA;  134722 MW;  BE7EEF2A5C28AA2D CRC64;

  Query Match             6.0%;  Score 563;  DB 1;  Length 1211;
  Best Local Similarity   21.2%;  Pred. No. 2.3e-27;
  Matches  178;  Conservative   88;  Mismatches  233;  Indels  342;  Gaps   26;

Qy         74 EHCNWEAWGEWSECSRTCGGGASYSLRRCL          CRNPLSQNTRYKTCNHECPPPA 127
              :||:| ||   :   |||||| |    :    |:|             || | |      :: ||   |||
Db        566 KDGSWGAWGPWGQCSRTCGTGVKFRTRQCDNPHPANGGRTCSGLAYDFQLCNRQDCPDSL 619

Qy        128 LDFRAQQCSAYN         DVQYQGHYYEWLPRYNDPAAPCALKCHAQGQNLVVELAPKV 180
               ||| :||    ::              | |        : |||    :  |         | | | ::       || :     |
Db        620 ADFREEQCRQWDLYFEHGDAQ       HHWLPHEHRDAKERCHLYCESRETGEVVSMKRMV 674

Qy        181 LDGTRCN  TDSLDMCISGICQAVGCDRQLGSNAKEDNCGVCAGDGSTCRLVRGQSKSHVS 239
               |||||:     |:     :|:   |  |:  |||:    :||:  :||   ||||  ||  |   |:::|:|          |
Db        675 HDGTRCSYKDAFSLCVRGDCRKVGCDGVIGSSKQEDKCGVCGGDNSHCKVVKGTFTR    S 732

Qy        240 FEKKEE NVIAVPLGSRSVRITVKGPAHLFIESKTLQGSK          GEHSFN SPGVFVVEN 293
              |:|           :    :|  | |:|  :  |                    :   | |:    |            |:    :  |       |:
Db        733 PKKHGYIKMFEIPAGARHLLIQEVDATSHHLAVKNLETGKFILNEENDVDASSKTFIAMG 792

Qy        294 TTVEPQRCSERQTFKIPGPLMAPFTFKTRYTAAKDSVVQFFFYQPISHQWRQTDFFPCTV 353
                    |::          |:|   :     |||
Db        793 VEWEYRDEDGRETLQTMGPL-------------------------------------- 812

Qy        354 TCGGGYQLNSAESVSTRSLRSVVPSHYCHYYPENVKPKPKLKECSMDPCPSSDGFKEIMPY 413
```

```
Db        813 ------------------------------------------------------------ 812

Qy        414 DHFQPLPRWEHNPWIACSVSCGGGIQRRSFVCVEESMHGEILQVEEWKCMYAPKPKVMQT 473
                         |   |   :   |       |   :   :  :|  :  |  |::         |::
Db        813 -----------HGTITVLVIPVG---DTRVSLTYKYMIHEDSLNVDD---------NNVLEE 851

Qy        474 CNLFDCPKWIAMEWSQCTVTCGRGLRYRVVLCINHRGEHVCGCNPQLKLHIKEECVIPIP 533
                :           :   |':
Db        852 DS-------VVYFWA--------------------------------------------- 859

Qy        534 CYKPKEKSPVEAKLPWLKQAQELEETRIATEEPTFIPEPWSACSTTCGPGVQVREVKCRV 593
                                  ||:                          || ||  || | |   :  ||
Db        860 ----- -----------LKK---------------------  WSPCSKPCGGGSQPTKYGCR 882

Qy        594 LLTFTQTETELPEEECEGPKLPTERPCLLEACDESPASRELDIPLPEDSETTYDWEYAGF 653
                                                         |  ||
Db        883 -------- ------------------------------RRLD------------------ 886

Qy        654 TPCTATCLGGHQFAIAVCLHIQTQQTVNDSLCDMVHRPPAMSQACNTEPC-PPRWHVGSW 712
                                          :  |:     |   :  :|  |:  :|||  :  |    |  |   |  |
Db        887   .          ...            HKMVHRGFCAALSKPKAIRRACNPQECSQPVWVTGEW 923

Qy        713 GPCSATCG-VGIQIRDVYCLHP  --GETPAPPEECRDEKPHALQACNQFDCPPGWHIEE 767
                |||  |||   |:|  |  |  |  |:  |                :  |  |  :|  :  :||:'   ||   |
Db        924 EPCSQTCGRTGMQVRSVRCIQPLHDNTTRSVHAKHCNDARPESRRACSRELCPGRWRAGP 983

Qy        768 WSQCSRTCGGGTQNRRVTCRQLLTDGSFLNLSDELCQGPKASSHKSCARTDCPPHLAVGD 827
                | |||  '|||  |:|  |  ,  ||        |  ||          :||   :    :   ::|        ||  :   :  |
Db        984 WSQCSVTCGNGTQERPVPCR--TADDSF-----GICQEERPETARTCRLGPCPRN  TSD 1034

Qy        828 WSKCSVSCGVGIQRRKQVCQRLAAKGRRIPLSEMMCR      ---LLPCPPLV  KSCQ 876
                ||  |   |        :                    ::::::||              :  ||                :||:   :    :||
Db       1035 PSEFTSVVQWLSEPTPLCTLEEINSFDRDQDLEE  CFRMLVLPKYCSTPCYNRLSCKSCN 1094

Qy        877 M 877
                :
Db       1095 L 1095


RESULT 10
AT10_MOUSE
ID   AT10_MOUSE     STANDARD;      PRT;   450 AA.
AC   P58459;
DT   15-JUN-2002 (Rel. 41, Created)
DT   15-JUN-2002 (Rel  41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   ADAMTS-10 (EC 3.4.24. ) (A disintegrin and metalloproteinase with
DE   thrombospondin motifs 10) (ADAM-TS 10) (ADAM-TS10) (Fragment).
GN   ADAMTS10.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Seldin M.F., Apte S.S.;
RT   "A novel gene of the ADAMTS family predicts ADAMTS-10, a
RT   metalloprotease with unique structural features and expression
RT   pattern.";
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).
CC    !- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
CC       matrix (By similarity).
CC   -!- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE-1 DOMAINS ARE IMPORTANT
CC       FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX (By
CC       similarity).
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12B.
CC   -!- SIMILARITY: CONTAINS AT LEAST 4 TSP TYPE 1 DOMAINS.
CC   --------------------------------------  ---------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -------------------------   -------------      ------------------    ---
DR   EMBL; AF302012; AAK97226.1; -.
DR   InterPro; IPR000884; TSP1.
DR   SMART, SM00209; TSP1; 4.
DR   PROSITE; PS50092; TSP1; 1.
KW   Hydrolase; Metalloprotease, Zinc, Glycoprotein; Repeat;
KW   Extracellular matrix.
FT   NON_TER        1      1
FT   DOMAIN        <1     51       CYS-RICH.
FT   DOMAIN        52    174       SPACER.
FT   DOMAIN       171    232       TSP TYPE-1 1.
FT   DOMAIN       234    290       TSP TYPE-1 2.
FT   DOMAIN       294    348       TSP TYPE-1 3.
FT   DOMAIN       353    399       TSP TYPE-1 4.
FT   CARBOHYD      86     86       N LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     141    141       N LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     238    238       N LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   450 AA,  48861 MW,  2377D0E2CFBFBFCA CRC64;

  Query Match             5.8%;  Score 544.5;  DB 1;  Length 450;
  Best Local Similarity   28.2%;  Pred. No. 9.4e-27;
  Matches  159;  Conservative   67;  Mismatches  186;  Indels  129;  Gaps   22;

Qy        177 APKVLLLCSCNLSELLMCLLLLQAVCTRCLGGNAKELNLGVCAGLLLTKLVKGQSKL 236
                |   |:|||  |    |::|:|:||  |:  ||||  |||:  :||  |  ||  ||||  |     :  |     |
Db          1 AAAVVDGTPCRPDTVDICVSGECKHVGCDRVLGSDLREDKCRVCGGDGSACETIEGVF-S 59

Qy        247 EVSPLPKKENVIAVPLGSRAVRITVKGPAHLI                 LPKTLQGRGE 280
                      |        |:|:  :|                     ||    |:||               ||    |:|    |
Db         60 PALPGTGYEDVVWIP          KGSVHIFIQDLNLSLSHLALKGDQESLLLEGLPGT 109

Qy        281 HSFNSTGVLVVENTIVEFQRGSER-QTFKIPGPLMADFTFKTPYIAAKDSVVQFFFYQPI 339
                        |        :    ||         ::|  ::   |:  :     ||:  |     |                |:        :   :   |     ||
Db        110 ---PQPHRLPLAGTTFHLRQGPDQAQSLEALGPINASLIIMV LAQAELPALHYRFNAPI 165

Qy        340 S        HQWKQTDFFPCTVTCGGGYQLNSAECVDIRLKRVVPDHYCHYYPENVKPKPKL 393
                :            :   |          :       |:      |     ||    |:          ||  :                 |   |||      :    :          ||  :
Db        166 ARDALPPYSWHYAPWTKCSAQCAGGSQVQVVECRNQLPSSAVAPHYCSGHSK  LPK RQ 222

Qy        394 KECSMDPCPSSDGFKEIMPYDHFQPLPRWEHNPWTACSVSCGGGIQRRSFVCVEESMHGE 453
                :  |:  :|||                            |  |         |:  ||  ||    |::  ||  ||                |
Db        223 RACNTPPCP  -                   PSWVVGNWSECGRSCLASVRSRSVVCQRKVSAAE 265

Qy        454 ILQVEEWKCMYAPKPKVMQTCNLFDCPKWIAMEWSQCTVTCGRGLRYRVVLCINHRGEH 512
                   :::    |       |:|  |::  |          |  |:|     ::||:||  :||  |||:|||||  :
Db        266 EKALDDSACP-QPRPPVLEACQGPMCPPEWATLDWSECTPSCGPGLRHRVVLCKS ---- 319

Qy        513 VGGCNPQLKLHIKEECVIP    IPCYKPKEKSPVEAKLPWLKQAQELEETRIATEEPTF 568
                             :      :|       :|     ||        |  :      |               :            |    |    |
Db        320 -----          ADQRSTLPPGHCLPAAKP--PSTMRCNL------RRCPPARWVTSE--- 357

Qy        569 IPEPWSACSTTCGPGVQVREVKCRVLLTFTQTETELPEEECEGPKLPTER  PPCLLEACD 626
                      |    |||  ||   |   |   |   |:|                  :  |     |     ||              ||    ||      ::  |:
Db        358 ----WGECSTQCGLGQQQRTVRC--------TSHTGQPSREC------TEALRPSTMQQCE 400

Qy        627 ESPASRELDIPLPEDSETTYDWEYAGFTP                       CTATCLG 662
                    |   :    :|   :   |     |         :  |                        |    ||  |
Db        401 ---AKCDSVVPPGDGPEECKDVNKVAYCPLVLKFQFCSRAYFRQMCCKTCQG 449


RESULT 11
ATS6_HUMAN
ID   ATS6_HUMAN     STANDARD;      PRT;   860 AA.
AC   Q9UKP5;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   ADAMTS-6 precursor (EC 3.4.24. ) (A disintegrin and metalloproteinase
DE   with thrombospondin motifs 6) (ADAM-TS 6) (ADAM-TS6).
GN   ADAMTS6.
```

```
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99395124; PubMed=10464288;
RA   Hurskainen T.L., Hirohata S., Seldin M.F., Apte S.S.;
RT   "ADAM-TS5, ADAM-TS6, and ADAM-TS7, novel members of a new family of
RT   zinc metalloproteases.";
RL   J. Biol. Chem. 274:25555-25563(1999).
CC   -!- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
CC       matrix (By similarity).
CC   -!- TISSUE SPECIFICITY: EXPRESSED AT LOW LEVELS IN PLACENTA AND BARELY
CC       DETECTABLE IN A NUMBER OF OTHER TISSUES.
CC   -!- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE-1 DOMAINS ARE IMPORTANT
CC       FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX.
CC   -!- PTM: THE PRECURSOR IS CLEAVED BY A FURIN ENDOPEPTIDASE (BY
CC       SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12B.
CC   -!- SIMILARITY: CONTAINS 1 DISINTEGRIN-LIKE DOMAIN.
CC   -!- SIMILARITY: CONTAINS 2 TSP TYPE-1 DOMAINS.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AF140674; AAD56357.1; -.
DR   HSSP; P34179; 1IAG.
DR   MEROPS; M12.230; -.
DR   Genew; HGNC:222; ADAMTS6.
DR   MIM; 605008; -.
DR   InterPro; IPR001762; Disintegrin.
DR   InterPro; IPR002870; Pep_M12B_propep.
DR   InterPro; IPR001590; Reprolysin.
DR   InterPro; IPR000884; TSP1.
DR   InterPro; IPR006026; Zn_MetPeptdse.
DR   Pfam; PF00090; tsp_1; 1.
DR   Pfam; PF01421; Reprolysin; 1.
DR   Pfam; PF01562; Pep_M12B_propep; 1.
DR   SMART; SM00209; TSP1; 1.
DR   PROSITE; PS50215; ADAM_MEPRO; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; 1.
DR   PROSITE; PS50092; TSP1; 1.
DR   PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
KW   Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
KW   Repeat; Extracellular matrix.
FT   SIGNAL        1     21       POTENTIAL.
FT   PROPEP       22    244       BY SIMILARITY.
FT   CHAIN       245    860       ADAMTS 6.
FT   METAL       403    403       ZINC (CATALYTIC) (BY SIMILARITY).
FT   ACT_SITE    404    404       BY SIMILARITY.
FT   METAL       407    407       ZINC (CATALYTIC) (BY SIMILARITY).
FT   METAL       413    413       ZINC (CATALYTIC) (BY SIMILARITY).
FT   DOMAIN      413    509       DISINTEGRIN-LIKE.
FT   DOMAIN      510    566       TSP TYPE-1 1.
FT   DOMAIN      567    668       CYS-RICH.
FT   DOMAIN      669    795       SPACER.
FT   DOMAIN      796    852       TSP TYPE-1 2.
FT   DOMAIN       68     71       POLY-ARG.
FT   DOMAIN      662    665       POLY-GLY.
FT   CARBOHYD     99     99       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    172    172       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    222    222       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    234    234       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    676    676       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    843    843       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   860 AA;  97098 MW;  E57213015DECB205 CRC64;

  Query Match             5.7%;  Score 529.5;  DB 1;  Length 860;
  Best Local Similarity   28.9%;  Pred. No. 1.8e-25;
  Matches  124;  Conservative   58;  Mismatches  136;  Indels  111;  Gaps   12;

Qy         75 DGNWDAWGDWSDCSRTCGGGASYSLRRCLT         GRNCEGQNLRYKTCNHDCPPDAE 128
              |  |    |   | :||||||| | ||| | :         |: | :  |:::|:    ||   :
Db        510 DGGWGPWSLWGECSRTCGGGVSSSLRHCDSPAPSGGGKYCLGERKRYRSCNTDPCPLGSR 569

Qy        129 DLRAQQCSAYNDVQYQGHYYEWLPRYNDPAAPCALKCHAQGQNLVVELAPKVLDGTRCNT 188
              :||  :||: ::::  ::|  ||  |  |            :|||  |  |:|  |     |  || :|||:||
Db        570 DFREKQCADFDNMPFRGKYYNWKPYTGGGVKPCALNCLAEGYNFYTERAPAVIDGTQCNA 629

Qy        189 DSLDMCISGICQAVGCDRQLGSNAKEDNCGVCAGDGSTCRLVRGQSKSHVSLEKREENVI 248
              |:||:|||:|  |:  ||:|   |||:|:|| |  || |  |||:  :  |       :  |       |:
Db        630 DSLDICINGECKHVGCDNILGSDAREDRCRVCGGDGSTCDAIEGFFNDSLPRGGYMEVV 688

Qy        249 AVPLGSRSVRITVKGPAHLFIESKTLQGS      KGEHSFNSPGVFVVENTTVEFQRGS E 303
               :| ||  : :       :  :|   |: :|          |  : : |   |  |   |     ::| : |
Db        689 QIPRGSVHIEVREVAMSKNYIALKS EGDDYYINGAWTIDWPRKFDVAGTAFHYKRPTDE 747

Qy        304 RQTFKTPGPLMADFIFKTRYTAAKDSVVQFFFYQPISHQWRQTDFFPCTVTCGGGYQLNS 363
               : : :   ||                      ::::      |   :     |        | |  |   |  ::
Db        748 PESLEALGP               ISENLIVMVLLQEQNLGIRYKFNVPITRTGSGDNEV     792

Qy        364 ASCVDIRLKRVVPDHYCHYYPENVKPKPKLKECSMDPCPSSDGFKEIMPYDHFQPLPRWE 423
                                                                ||                 |
Db        793 ------------------------------------------------GF                 TWN 797

Qy        424 HNPWTACSVSCGGGIQ            LQSRTT          VCVEESMHGEILQVEEWKCMYAPK 467
              |  ||: ||  :| ||              ||    :             :|:  :  :  |  |           |  :|
Db        798 HQPWSECSATCAGGKMPTRQPTQRARWRTKHILSYALCLLKKLIGNI         SCREA    849

Qy        468 PKVMQTCNL 476
                   :|||
Db        850     SSCNL 854


RESULT 12
ATS8_HUMAN
ID   ATS8_HUMAN     STANDARD;      PRT;   890 AA.
AC   Q9UP79; Q9NZS0;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   ADAMTS 8 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase
DE   with thrombospondin motifs 8) (ADAM-TS 8) (ADAM-TS8) (METH-2)
DE   (METH-8).
GN   ADAMTS8 OR METH2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Lung;
RX   MEDLINE=99367466; PubMed=10438512;
RA   Vazquez F., Hastings G., Ortega M.A., Lane T.F., Oikemus S.,
RA   Lombardo M., Iruela-Arispe M.L.;
RT   "METH-1, a human ortholog of ADAMTS-1, and METH-2 are members of a new
RT   family of proteins with angio-inhibitory activity.";
RL   J. Biol. Chem. 274:23349-23357(1999).
RN   [2]
RP   SEQUENCE OF 195-440 FROM N.A.
RX   MEDLINE=20079168; PubMed=10610729;
RA   Georgiadis K.E., Hirohata S., Seldin M.F., Apte S.S.;
RT   "ADAM-TS8, a novel metalloprotease of the ADAM-TS family located on
RT   mouse chromosome 9 and human chromosome 11.";
RL   Genomics 62:312-315(1999).
CC   -!- FUNCTION: HAS ANTI-ANGIOGENIC PROPERTIES.
CC   -!- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
CC       matrix (By similarity).
CC   -!- TISSUE SPECIFICITY: HIGLY EXPRESSED IN ADULT AND FETAL LUNG, LOWER
CC       EXPRESSION IN BRAIN, PLACENTA, HEART, STOMACH AND FETAL BRAIN AND
CC       KIDNEY.
CC   -!- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE-1 DOMAINS ARE IMPORTANT
CC       FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX.
CC   -!- PTM: THE PRECURSOR IS CLEAVED BY A FURIN ENDOPEPTIDASE (BY
CC       SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12B.
CC   -!- SIMILARITY: CONTAINS 1 DISINTEGRIN-LIKE DOMAIN.
CC   -!- SIMILARITY: CONTAINS 2 TSP TYPE-1 DOMAINS.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AF060153; AAD48081.1; -.
DR   EMBL; AF175283; AAF25806.1; -.
DR   HSSP; P34179; 1IAG.
DR   MEROPS; M12.226; -.
DR   Genew; HGNC:224; ADAMTS8.
DR   MIM; 605175; -.
DR   InterPro; IPR001762; Disintegrin.
DR   InterPro; IPR002870; Pep_M12B_propep.
DR   InterPro; IPR001590; Reprolysin.
DR   InterPro; IPR000884; TSP1.
DR   InterPro; IPR000130; Zn_MIpeptdse.
DR   Pfam; PF00090; tsp_1; 2.
DR   Pfam; PF01421; Reprolysin; 1.
DR   Pfam; PF01562; Pep_M12B_propep; 1.
DR   SMART; SM00209; TSP1; 2.
DR   PROSITE; PS50215; ADAM_MEPRO; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; 1.
DR   PROSITE; PS50092; TSP1; 1.
DR   PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
KW   Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
KW   Repeat; Extracellular matrix; Heparin-binding.
FT   SIGNAL        1     27       POTENTIAL.
FT   PROPEP       28    214       BY SIMILARITY.
FT   CHAIN       215    890       ADAMTS-8.
FT   METAL       364    364       ZINC (CATALYTIC) (BY SIMILARITY).
FT   ACT_SITE    365    365       BY SIMILARITY.
FT   METAL       368    368       ZINC (CATALYTIC) (BY SIMILARITY).
FT   METAL       374    374       ZINC (CATALYTIC) (BY SIMILARITY).
FT   DOMAIN      439    526       DISINTEGRIN-LIKE.
FT   DOMAIN      527    583       TSP TYPE-1 1.
FT   DOMAIN      584    690       CYS-RICH.
FT   DOMAIN      691    832       SPACER.
FT   DOMAIN      833    890       TSP TYPE-1 2.
FT   DOMAIN      202    205       POLY-PRO.
FT   CARBOHYD    345    345       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    401    401       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    466    466       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    491    491       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    600    600       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT    195    195       E -> R (IN REF. 2).
FT   CONFLICT    413    440       YLTELLDGGHGDCLLDAPGAALPLPTGL -> FSGCHLQGW
FT                                IHFKYLCKCVSELKCDLMP (IN REF. 2).
SQ   SEQUENCE   890 AA;  96671 MW;  57D70EE03D5739D3 CRC64;

  Query Match             5.[illegible]%;  Score 496.5;  DB 1;  Length 890;
  Best Local Similarity   31.9%;  Pred. No. 2.1e-23;
  Matches  122;  Conservative   55;  Mismatches  131;  Indels   75;  Gaps   19;

Qy         75 DGNWDAWGDWSDCSRTCGGGASYSLRRCLT------GRNCEGQNIRYKTCSNHDCPPDAE 128
              || |  || | :|||:||||  :| | |         || | |:  :|::|   :|||| :
Db        527 DGGWAPWGPWGECSRTCGGGVQFSHRECKDPEPQNGGRYCLGRRAKYQSCHTEECPPDGK 586

Qy        129 DFRAQQCSAYNDVQY---QGHYYEWLPRYN--DPAAPCALKCHAQGQNLVVELAPKVLDG 183
               || |||  ||    |     :  :|:|:|    |    | | | |:|::        ||:||
Db        587 SFREQQCEKYNAYNYTDMDGNLLQWVPKYAGVSPRDRCKLFCRARGRSEFKVFEAKVIDG 646

Qy        184 TRGNTPSLDMCISGICQAVGCDRQLGSNAKEDNCGVCAGDGSTCRLVRGQSKSHVSPEKR 243
               | |   ::| :|: | |    |||   : |    | | |||| | |::|| | |      ::|
Db        647 ILGGEELLAIVKIQGVEA[illegible]NY 702

Qy        244 LEN VIAVPLGSRSVRIIVKGHAHL[illegible]RGVEVV 291
                  | :: :| |   :   | ||   :|   :::          |: : |:: |           :|
Db        703 GYNDIVIIPAG--AINIDVKQRSHPGVQNDGNYLALKTADGQYLLNGNLAISAIEQDILV 760

Qy        292 ENTTVEFQRGS[illegible]KTEYTAAEQVVQIF 334
              : | :::  ||     || |:|: :| ||       :           | :||     : | |
Db        761 FGTILKYS-GSIATLERIQSF[illegible] 819

Qy        335 ---------FYQPISH QWRQTDFFPCTVTCGGGYQLNSAECVDIRLKRVVPDHY----CH 381
                           ||: | ||   |:   |: ||| |:|   : || |          |       |:
Db        820 SKEPATTNIIQPLLHAQWVLGDWSECSSTCGAGWQRRTVECRD-------PSGQASATCN 872

Qy        382 YYPENVKPKPKLKECSMDPCP 402
                 ||: ||         |     ||
Db        873 KALKPEDAKP------CESQLCP 889


RESULT 13
ATS8_MOUSE
ID   ATS8_MOUSE     STANDARD;      PRT;   905 AA.
AC   P57110;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   ADAMTS-8 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase
DE   with thrombospondin motifs 8) (ADAM-TS 8) (ADAM-TS8) (METH-2).
GN   ADAMTS8.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20079168; PubMed=10610729;
RA   Georgiadis K.E., Hirohata S., Seldin M.F., Apte S.S.;
RT   "ADAM-TS8, a novel metalloprotease of the ADAM-TS family located on
RT   mouse chromosome 9 and human chromosome 11.";
RL   Genomics 62:312-315(1999).
CC   -!- FUNCTION: HAS ANTI-ANGIOGENIC PROPERTIES (BY SIMILARITY).
CC   -!- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
CC       matrix (By similarity).
CC   -!- TISSUE SPECIFICITY: EXPRESSED SPECIFICALLY IN ADULT LUNG AND HEART
CC       AND LOW EXPRESSION DURING MOUSE DEVELOPMENT.
CC   -!- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE-1 DOMAINS ARE IMPORTANT
CC       FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX.
CC   -!- PTM: THE PRECURSOR IS CLEAVED BY A FURIN ENDOPEPTIDASE (BY
CC       SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12B.
CC   -!- SIMILARITY: CONTAINS 1 DISINTEGRIN-LIKE DOMAIN.
CC   -!- SIMILARITY: CONTAINS 2 TSP TYPE-1 DOMAINS.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AF175282; AAF25805.1; -.
DR   HSSP; P34179; 1IAG.
DR   MEROPS; M12.226; -.
DR   MGD; MGI:1353468; Adamts8.
DR   InterPro; IPR001762; Disintegrin.
DR   InterPro; IPR001590; Reprolysin.
DR   InterPro; IPR000884; TSP1.
DR   InterPro; IPR000130; Zn MIpeptdse.
DR   Pfam; PF00090; tsp_1; 2.
DR   Pfam; PF01421; Reprolysin; 1.
DR   SMART; SM00209; TSP1; 2.
DR   PROSITE; PS00142; ZINC PROTEASE; 1.
DR   PROSITE; PS50215; ADAM MEPRO; 1
DR   PROSITE; PS50092; TSP1; 1.
DR   PROSITE; PS00427; DISINTEGRIN 1; FALSE NEG.
KW   Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
KW   Repeat; Extracellular matrix; Heparin binding.
FT   SIGNAL          1     28        POTENTIAL.
FT   PROPEP         29    228        BY SIMILARITY.
FT   CHAIN         229    905        ADAMTS 8.
FT   METAL         378    378        ZINC (CATALYTIC) (BY SIMILARITY).
FT   ACT SITE      379    379        BY SIMILARITY.
FT   METAL         382    382        ZINC (CATALYTIC) (BY SIMILARITY).
FT   METAL         388    388        ZINC (CATALYTIC) (BY SIMILARITY).
FT   DOMAIN        453    541        DISINTEGRIN LIKE.
FT   DOMAIN        542    598        TSP TYPE 1 1.
FT   DOMAIN        599    705        CYS RICH.
FT   DOMAIN        706    847        SPACER.
FT   DOMAIN        848    905        TSP TYPE 1 2.
FT   CARBOHYD      415    415        N LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD      480    480        N LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD      506    506        N LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD      615    615        N LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   905 AA;  98879 MW;  12404132B33A0CAE CRC64;

  Query Match             12.6%;  Score 484.5;  DB 1;  Length 905;
  Best Local Similarity   27.4%;  Pred. No. 1.2e-22;
  Matches  128;  Conservative   64;  Mismatches  153;  Indels  113;  Gaps   19;

Qy         34 FLPEPALSPQGSPLEPLLTEQFLLYRYDDQTSRNIRS                       70
              | |:|    |  | :||    :     :|  |:    :|::
Db        471 FGPDERHCPNIS VEDEVQQCARHRDSDEPTCHIKNGSLLWADGTPCGPGHDCLDGSCV 529

Qy         71 DED          KDCNNGAWTHWGSWSQCSRTCGGGVQFAVRQCD     GRNCEGQNIEYK 116
                ||           |::|   .   |   |:|||:|:   :| | |               | |  :: :: :
Db        530 TKEDVENPKAVVPGNWGPWEEWGQCSRTCGGGIQFGNRECNNPMPQNGGHPCSGERVKYQ 589

Qy        117 TCSNEDCPPDAESFRAQQC    SAYNEVQYQGHYYEWLLKYN  DRAAPCALKCHAQGQN 171
                 :|:   :|||:  :  ||  |||      :|:|           |::  :|:|:|:       |    |  |  |  |:|::
Db        590 SCNTEECPPNGKSFREQQCEKYNAYNHTDLDGNFLQWVPKYSGVSPRDRCKLFCRARGRS 649

Qy        172 LVVELAPKVLDGTRCNTDSLDMCISGICQAVGCDRQLGSNAKEDNCGVCAGDGSTCRLVR 231
                          ||:|||  |   |:  :|:  |  |      |||   :  |     |  |  ||:|  |  :  ||  :
Db        650 EFKVFEAKVIDGTLCGPETLAICVRGQCVKAGCDHVVNSPKKLDKCGVCGGKGTACRKIS 709

Qy        232 GQSKSHVSPEKREENVIAVPLGSRSVRITVKGPAHLFIESK     TLQGSKGEHSFNS   285
              |    |              ::: :|  :   :   | ||   :|   : :        |:  :  ::: |
Db        710 G    SFTPFSYGYNDIVTIPAG  ATNIDVKQRSHPGVKNDGNYLALKTANGQYLLNGNL 764

Qy        286          PGVFVVENTTVEFQRGS     EE QTFK IPGPLMADFIF            KIRYT     324
                          :|: |  :::   :.        || |:|: :|  ||        :             : |
Db        765 AISAIEQDILVKGTILKYSGSMATLERLQSFQALPEPLTVQLLTVSGEVFPPKVRYTFF 823

Qy        325                  AAKGSVVQEEEYQPTSHQWRQTDFFPSTIVTGGGTYQENSAD            366
                                  :|:                 |  :|    |:  |    ||  |  :|    :  :|
Db        824 VPNDMDFSVQNSKERATTNIIQSLPSAEWVLGDWSECPSTCRGSWQRRTVECRDPSGQAS 883

Qy        367     VDTRLKRVVPDHYCHYYPENVKPKPKLKECSMDPCP 402
                    |   ||                  ||: ||            |     ||||
Db        884 DTCDEALK                  PEDAKP          CGSQPCP 904


RESULT 14
ATS5_MOUSE
ID   ATS5_MOUSE     STANDARD;      PRT;   930 AA.
AC   Q9R001;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   ADAMTS 5 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase
DE   with thrombospondin motifs 5) (ADAM-TS 5) (ADAM-TS5) (Aggrecanase 2)
DE   (ADMP-2) (Implantin).
GN   ADAMTS5.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99395124; PubMed=10464288;
RA   Hurskainen T.L., Hirohata S., Seldin M.F., Apte S.S.;
RT   "ADAM-TS5, ADAM-TS6, and ADAM-TS7, novel members of a new family of
RT   zinc metalloproteases.";
RL   J. Biol. Chem. 274:25555-25563(1999).
CC   -!- FUNCTION: CLEAVES AGGRECAN, A CARTILAGE PROTEOGLYCAN, AND MAY BE
CC       INVOLVED IN ITS TURNOVER. MAY PLAY AN IMPORTANT ROLE IN THE
CC       DESTRUCTION OF AGGRECAN IN ARTHRITIC DISEASES. MAY PLAY A ROLE IN
CC       PROTEOLYTIC PROCESSING MOSTLY DURING THE PERI-IMPLANTATION PERIOD.
CC   -!- CATALYTIC ACTIVITY: Cleaves aggrecan at the 392-Glu-|-Ala-393
CC       site.
CC   -!- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
CC       matrix (By similarity).
CC   -!- DEVELOPMENTAL STAGE: EXPRESSED SPECIFICALLY IN THE PERI-
CC       IMPLANTATION PERIOD IN EMBRYO AND TROPHOBLAST AND AT LOW OR
CC       UNDECTABLE LEVEL THEREAFTER.
CC   -!- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE-1 DOMAINS ARE IMPORTANT
CC       FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX.
CC   -!- PTM: THE PRECURSOR IS CLEAVED BY A FURIN ENDOPEPTIDASE (BY
CC       SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12B.
CC   -!- SIMILARITY: CONTAINS 1 DISINTEGRIN-LIKE DOMAIN.
CC   -!- SIMILARITY: CONTAINS 2 TSP TYPE-1 DOMAINS.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AF140673; AAD56356.1; -.
DR   MEROPS; M12.225; -.
DR   MGI; MGI:1346321; Adamts5.
DR   InterPro; IPR001762; Disintegrin.
DR   InterPro; IPR002870; Pep_M12B_propep.
DR   InterPro; IPR001590; Reprolysin.
DR   InterPro; IPR000884; TSP1.
DR   InterPro; IPR000130; Zn_MIpeptdse.
DR   Pfam; PF00090; tsp_1; 2.
DR   Pfam; PF01421; Reprolysin; 1.
DR   Pfam; PF01562; Pep_M12B_propep; 1.
DR   SMART; SM00209; TSP1; 2.
DR   PROSITE; PS50215; ADAM_MEPRO; 1
DR   PROSITE; PS50092; TSP1; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; 1.
DR   PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
KW   Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
KW   Repeat; Extracellular matrix.
FT   SIGNAL          1     21        POTENTIAL.
FT   PROPEP         22    261        POTENTIAL.
FT   CHAIN         262    930        ADAMTS 5.
FT   SITE          209    209        CYSTEINE SWITCH (POTENTIAL)
FT   METAL         410    410        ZINC (CATALYTIC) (BY SIMILARITY).
FT   ACT SITE      411    411        BY SIMILARITY.
FT   METAL         414    414        ZINC (CATALYTIC) (BY SIMILARITY).
FT   METAL         420    420        ZINC (CATALYTIC) (BY SIMILARITY)
FT   DOMAIN        485    566        DISINTEGRIN LIKE.
FT   DOMAIN        567    623        TSP TYPE 1 1.
FT   DOMAIN        624    731        CYS RICH.
FT   DOMAIN        732    874        SPACER.
FT   DOMAIN        875    930        TSP TYPE 1 2
FT   DOMAIN       41     46       POLY-ALA.
FT   DOMAIN      257    261       POLY-ARG.
FT   CARBOHYD    498    498       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    728    728       N LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    802    802       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    807    807       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   930 AA;  101780 MW;  84DF84B26170D4DC CRC64;

  Query Match             5.2%;  Score 482;  DB 1;  Length 930;
  Best Local Similarity   26.3%;  Pred. No. 1.8e 22;
  Matches  133;  Conservative   67;  Mismatches  170;  Indels  136;  Gaps   21;

Qy         34 FLPEFALSP----------------QCSF----------LEDT---TCEQFLTYRYDDQTSRN 67
              | ||::: ;                ||             :| |       |    |   :  |:| :
Db        501 FCPEYSVCPGMDVCARLWCAVVRQGQMVCITKKLPAVEGTPCGKGRVCLQGKCVDKTKKK 560

Qy         68 TRSDEDKDRNWEAWCPWSECSRTCGGGASYSLRRC         -LICRNCEGQNIRYKTCSNH 121
                |       ||| :|| |    |||:||||   :: | |          :|| | |:     |::||
Db        561 YYS-TSSHGNWGSWGPWGQCSRSCGGGVQFAYRHCNNPAPRNSGRYCTGKRAIYRSCSVT 619

Qy        122 DCPPDAHDFRAQQCSAYNDVQYQGH     -YYEWLPRYND  PAAPCALKCHAQGQNLVVE 175
               |||: : || :|| | |   |            : ||:|:|       ||   | | | |:|       |
Db        620 PCPPNGKSFRHEQCEAKNGYQSDAKGVKTFVEWVPKYAGVLPADVCKLTCRAKGTGYYVV 679

Qy        176 LAPKVLDGTRCNTDSLDMCISGICQAVGCDRQLGSNAKEDNCGVCAGDGSTCRLVRG    232
               :||| ||| |      |  :|: | |     |||  :||   : | |||| || |:|   : |
Db        680 FSPKVTDGTECRPYSNSVCVRGRCVRTGCDGIIGSKLQYDKCGVCGGDNSSCTKIIGTFN 739

Qy        233 -QSKSHVSPEKREENVIAVPLGSRSVRI--      -IVKGPAHLFIESKTLQGSKGEHSFN 284
               :||  :          :|: :| ||:   :::              : ||:| :: ||       ||:   |
Db        740 KKSKGYT-------DVVRIPEGATHIKVRQFKAKDQTRFTAYLALKKK---  GEYLIN 787

Qy        285       ---SPGVFVVENITVEFQRGSERQTFKIPGPLMADFIFKIRYIAAKD-SVVQFFFY 336
                     |  :   :   |  :  :      ||             ||:      |:|  |:   :||
Db        788 GKYMISTSETIIDINGTVMNYSGWSHRD----------DFLHGMGYSATKEILIVQILAT 837

Qy        337 QPISHQWRQTDFFPCTVTCGGGYQLNSAECVDIRLKRVVPDHYCHYYPENVKPKPKLKEC 396
               |                                   : : :|            |   :  |:     |
Db        838 DP----------------------------TKALGVR---------YSFFVPKKITQKVN---- 860

Qy        397 SMDPCPSSDGFKEIMPYDHFQPLPRWEHNPWTACSVSCGGGIQRRSFVCVE---ESMHGE 453
                      ; |   :: |    |    | :|    || |||  :|    |     |:   |  :     :     |
Db        861 ---SVISHGSNKVGP--RSTQL-QWVTGPWLACSRTCGTGWHTRTVQCQDGNRKLAKG 912

Qy        454 IIQVEEWKCMYAPKPKVMQTCNLFDC 479
                          |: : :,    : | |  |
Db        913 --------CLLSQRPSAFKQCLLKKC 930


RESULT 15
ATS5_HUMAN
ID   ATS5_HUMAN     STANDARD;      PRT;   930 AA.
AC   Q9UNA0; Q9UKP2;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   ADAMTS-5 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase
DE   with thrombospondin motifs 5) (ADAM-TS 5) (ADAM-TS5) (Aggrecanase-2)
DE   (ADMP-2) (ADAM-TS 11).
GN   ADAMTS5 OR ADMP2 OR ADAMTS11.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=99367476; PubMed=10438522;
RA   Abbaszade I., Liu R.-Q., Yang F., Rosenfeld S.A., Ross O.H.,
RA   Link J.R., Ellis D.M., Tortorella M.D., Pratta M.A., Hollis J.M.,
RA   Wynn R., Duke J.L., George H.J., Hillman M.C. Jr., Murphy K.,
RA   Wiswall B.H., Copeland R.A., Decicco C.P., Bruckner R., Nagase H.,
RA   Itoh Y., Newton R.C., Magolda R.L., Trzaskos J.M., Hollis G.F.,
RA   Arner E.C., Burn T.C.;
RT   "Cloning and characterization of ADAMTS11, an aggrecanase from the
RT   ADAMTS family.";
RL   J. Biol. Chem. 274:23443-23450(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20289733; PubMed=10830953;
RA   Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,
RA   Park H.-S., Toyoda A., Ishii K., Totoki Y., Choi D.-K., Soeda E.,
RA   Ohki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K., Polley A.,
RA   Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,
RA   Reichwald K., Rump A., Schillhabel M., Schudy A., Zimmermann W.,
RA   Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S.,
RA   Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E.,
RA   Minoshima S., Shimizu N., Nordsiek G., Hornischer K., Brandt P.,
RA   Scharfe M., Schoen O., Desario A., Reichelt J., Kauer G., Bloecker H.,
RA   Ramser J., Beck A., Klages S., Hennig S., Riesselmann L., Dagand E.,
RA   Wehrmeyer S., Borzym K., Gardiner K., Nizetic D., Francis F.,
RA   Lehrach H., Reinhardt R., Yaspo M.-L.;
RT   "The DNA sequence of human chromosome 21.";
RL   Nature 405:311-319(2000).
RN   [3]
RP   SEQUENCE OF 413-930 FROM N.A.
RC   TISSUE=Fetal brain;
RX   MEDLINE=99395124; PubMed=10464288;
RA   Hurskainen T.L., Hirohata S., Seldin M.F., Apte S.S.;
RT   "ADAM-TS5, ADAM-TS6, and ADAM-TS7, novel members of a new family of
RT   zinc metalloproteases.";
RL   J. Biol. Chem. 274:25555-25563(1999).
CC   -!- FUNCTION: CLEAVES AGGRECAN, A CARTILAGE PROTEOGLYCAN, AND MAY BE
CC       INVOLVED IN ITS TURNOVER. MAY PLAY AN IMPORTANT ROLE IN THE
CC       DESTRUCTION OF AGGRECAN IN ARTHRITIC DISEASES. MAY PLAY A ROLE IN
CC       PROTEOLYTIC PROCESSING MOSTLY DURING THE PERI-IMPLANTATION PERIOD.
CC   -!- CATALYTIC ACTIVITY: Cleaves aggrecan at the 392-Glu-|-Ala-393
CC       site.
CC   -!- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
CC       matrix (By similarity).
CC   -!- TISSUE SPECIFICITY: EXPRESSED AT LOW LEVEL, PRIMARILY IN PLACENTA
CC       BUT ALSO IN OTHER TISSUES, SUCH AS HEART AND BRAIN, AND ALSO
CC       CERVIX, UTERUS, BLADDER, ESOPHAGUS, RIB CARTILAGE,
CC       CHONDROBLASTOMA, FIBROUS TISSUE AND JOINT CAPSULE FROM AN
CC       ARTHRITIC PATIENT.
CC   -!- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE-I DOMAINS ARE IMPORTANT
CC       FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX.
CC   -!- PTM: THE PRECURSOR IS CLEAVED BY A FURIN ENDOPEPTIDASE (BY
CC       SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12B.
CC   -!- SIMILARITY: CONTAINS 1 DISINTEGRIN-LIKE DOMAIN.
CC   -!- SIMILARITY: CONTAINS 2 TSP TYPE-1 DOMAINS.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AF142099; AAD49577.1; -.
DR   EMBL; AP001698; BAA95504.1; -.
DR   EMBL; AP001697; BAA95503.1; -.
DR   EMBL; AF141293; AAF02493.1; -.
DR   HSSP; Q9PW35; 1BUD.
DR   MEROPS; M12.225; -.
DR   Genew; HGNC:221; ADAMTS5.
DR   MIM; 605007; -.
DR   InterPro; IPR001762; Disintegrin.
DR   InterPro; IPR002870; Pep_M12B_propep.
DR   InterPro; IPR001590; Reprolysin.
DR   InterPro; IPR000884; TSP1.
DR   InterPro; IPR000130; Zn_MTpeptdse.
DR   Pfam; PF00090; tsp_1; 2.
DR   Pfam; PF01421; Reprolysin; 1.
DR   Pfam; PF01562; Pep_M12B_propep; 1.
DR   SMART; SM00209; TSP1; 2.
DR   PROSITE; PS50215; ADAM_MEPRO; 1.
DR   PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
DR   PROSITE; PS50092; TSP1; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; 1.
KW   Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
KW   Repeat; Extracellular matrix.
FT   SIGNAL        1     16       POTENTIAL.
FT   PROPEP       17    261       POTENTIAL.
FT   CHAIN       262    930       ADAMTS-5.
FT   SITE        259    259       CYSTEINE SWITCH (POTENTIAL).
FT   METAL       410    410       ZINC (CATALYTIC) (BY SIMILARITY).
FT   ACT_SITE    411    411       BY SIMILARITY.
FT   METAL       414    414       ZINC (CATALYTIC) (BY SIMILARITY).
FT   METAL       420    420       ZINC (CATALYTIC) (BY SIMILARITY).
FT   DOMAIN      485    566       DISINTEGRIN-LIKE.
FT   DOMAIN      567    623       TSP TYPE-1 1.
FT   DOMAIN      624    731       CYS-RICH.
FT   DOMAIN      732    874       SPACER.
FT   DOMAIN      875    930       TSP TYPE-1 2.
FT   DOMAIN       37     41       POLY-ALA.
FT   DOMAIN      257    261       POLY-ARG.
FT   CARBOHYD    498    498       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    728    728       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    802    802       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    807    807       N-LINKED (GLCNAC. . .) (POTENTIAL)
FT   CONFLICT    138    138       A -> G (IN REF. 2)
FT   CONFLICT    614    614       R -> H (IN REF. 3).
FT   CONFLICT    692    692       P -> L (IN REF. 2).
SQ   SEQUENCE   930 AA;  101715 MW;  B64281502F28193B CRC64;

  Query Match              5.1%;  Score 477;  DB 1;  Length 930;
  Best Local Similarity   27.0%;  Pred. No. 3.8e-22;
  Matches  122;  Conservative   63;  Mismatches  157;  Indels  110;  Gaps   18;

Qy         62 DQTSENTRSDEDKECNWDAWGDWSDCSRTCGGGASYSLRRC  -     LDGRNCEGQNIRY 115
              |:|  :    :       |.|  :|   |   |||:|||   :: |  |            :| | |.:   :
Db        555 DEIKKKYYSLSSHCNWSSWSSWSQCSRSCGGGVQTAYRHCNNPAPRNNGRYCTGKRAIY 613

Qy        116 KLCNHEPCPPDAEDFRAQQCSAYNEVQYQGH      YYEWLPRYND  PAAPCALNCHAQG 169
              ::||    |||: : || :||  |  |  |         : ||:|:|    ||  . | | : :
Db        614 RSCSLMPCPPNGKSFRHEQCEAKNGYQSDAKGVKTFVEWVPKYAGVLPADVCKLTCRAKG 673

Qy        170 QNLVVELAPKVLDGTRCNTDSLDMCISGICQAVGCDRQLGSNAKEDNCGVCAGDGSTCRL 229
                      |    :|||  |||  |      |     :|:  | |      |||    :||    :  |  ||||  || |:
Db        674 TGYYVVFSPKVTDGTECRLYSNSVCVRGKCVRTGCDGIIGSKLQYDKCGVCGGDNSSCTK 733

Qy        230 VRG      QGKSHVSPEKREENVIAVPLGSRSVRI  -     IVKGPAHLFIESKTLQGSK 278
              . |      :||  :            :|: :| |: :::              :   |:| ::  |
Db        734 IVGTFNKKSKGYT          DVVRIPEGATHIKVRQFKAKDQTRFTAYLALKK       N 781

Qy        279 GEHSFN          SPGVFVVENTTVEFQRGSERQTFKIPGPLMADFIFKTRYTAAKDSV 330
              ||:   |          .   .   :   .   :   :   ||                ||:      |:      :      :
Db        782 GEYLINGKYMISTSETIIDINGTVMNYSGWSHRD           DFLHGMGYSATKEIL 831

Qy        331 VQFFFYQPISHQWRQTDFFPCTVTCGGGYQLNSAECVDIRLKRVVPDHYCHYYPENVKPK 390
              ||          |                                       : :|:|                   |     :     :     |
Db        832 VQILATDP                                       TKPLDVR            YSFFVPK    KST 856

Qy        391 PKLKECSMDPCPSSDGFKEIMPYDHFQPLPRWEHNPWTACSVSCGGGIQRRSFVCVE      447
              ||:                :|  |   ::   .           |:|      ||  |||  :|     .      |:    |   :
Db        857 PKVNSV          TSHGSNKVGSH      TSQPQWVTGPWLACSRTCDTGWHTRTVQCQDGNR 907

Qy        448 ESMHGEIQQVEEWKCMYAPKPKVMQTCNLFDC 479
               :                       .           :  |   .   |
Db        908 KLAKG                  CPLSQRPSAFKQCLLKKC 930

Search completed: March 22, 2003, 13:29:17
Job time : 45 secs
```

```
                    GenCore version 5.1.4 p5 4578
               Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on:          March 22, 2003, 13:09:44 ; Search time 70 Seconds
                                          (without alignments)
                                          4977.510 Million cell updates/sec

Title:           US-09-784-358-2
Perfect score:   9346
Sequence:        1 MACWICPWWVLLMVFMHGP..........LNLCGLDKYKQRCCQSCQEG 1691

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters:        671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :        SPTREMBL_21:*
                 1:   sp_archea:*
                 2:   sp_bacteria:*
                 3:   sp_fungi:*
                 4:   sp_human:*
                 5:   sp_invertebrate:*
                 6:   sp_mammal:*
                 7:   sp_mhc:*
                 8:   sp_organelle:*
                 9:   sp_phage:*
                 10:   sp_plant:*
                 11:   sp_rodent:*
                 12:   sp_virus:*
                 13:   sp_vertebrate:*
                 14:   sp_unclassified:*
                 15:   sp_rvirus:*
                 16:   sp_bacteriap:*
                 17:   sp_archeap:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                              SUMMARIES
                      %
Result              Query
   No.   Score  Match Length DB   ID                             Description

    1     5599    59.9   1023  4  Q9UL17                         Q9ul17 homo sapien
    2     4322    46.2    766  4  P82987                         P82987 homo sapien
    3     3729    39.9    761  6  Q95LQ2                         Q95lq2 macaca fasc
    4     2048    21.9    525  4  Q96RW4                         Q96rw4 homo sapien
    5     1483    15.9    417  4  Q9BXY3                         Q9bxy3 homo sapien
    6   1270.5    13.6   1914  5  Q95R33                         Q95r33 drosophila
    7   1183.5    12.7   2165  5  Q19791                         Q19791 caenorhabdi
    8   1110.5    11.9   1059  5  P90884                         P90884 caenorhabdi
    9    958.5    10.3   1280 11  Q9EPX2                          Q9epx2 mus musculu
   10    913.5     9.8   1235  4  O95428                         O95428 homo sapien
   11    898.5     9.6   2174  5  Q9GQR0                         Q9gqr0 drosophila
   12      891     9.5   3060  5  Q9VAV4                         Q9vav4 drosophila
   13      863     9.2   2167  5  O76840                         O76840 caenorhabdi
   14    860.5     9.2   3108  5  Q9U8C8                         Q9u8c8 manduca sex
   15    825.5     8.8   1572  5  O44938                         O44938 haemonchus
   16    774.5     8.3   1427  4  Q96L37                         Q96l37 homo sapien
   17    773.5     8.3    858  5  Q9V9Q7                         Q9v9q7 drosophila
   18    744.5     8.0   1207  4  Q8IE59                         Q8ie59 homo sapien
   19      720     7.7    132 11  Q9CX59                          Q9cx59 mus musculu
   20      717     7.7   1005  4  Q8IE56                         Q8ie56 homo sapien
   21      700     7.5    951  4  O60345                         O60345 homo sapien
   22    670.5     7.2   1072  4  Q8IE57                         Q8ie57 homo sapien
   23    659.5     7.1   1081  4  Q8IE60                         Q8ie60 homo sapien
   24      596     6.4   1159  4  Q8IEY8                         Q8iey8 homo sapien
   25      596     6.4   1223  4  Q8WXS8                         Q8wxs8 homo sapien
   26      593     6.3    898  4  Q9UFZ4                         Q9ufz4 homo sapien
   27      571     6.1   1223  4  Q8IE55                         Q8ie55 homo sapien
   28      571     6.1   1688  5  Q8SXB0                         Q8sxb0 drosophila
   29    570.5     6.1    950  4  Q8IE58                         Q8ie58 homo sapien
   30      558     6.0   1054  5  Q9W493                         Q9w493 drosophila
   31    543.5     5.8    622  4  Q9H8X0                         Q9h8x0 homo sapien
   32    532.5     5.7   1229  5  Q9VF61                         Q9vf61 drosophila
   33    499.5     5.3   5146  6  Q8SPM4                         Q8spm4 bos taurus
   34      499     5.3    538  4  Q9H8E4                         Q9h8e4 homo sapien
   35      490     5.2   1444  5  Q17591                         Q17591 caenorhabdi
   36    438.5     4.7   1290  4  Q9UPZ6                         Q9upz6 homo sapien
   37    438.5     4.7   2673  4  Q96SC3                         Q96sc3 homo sapien
   38    438.5     4.7   5636  4  Q96RW7                         Q96rw7 homo sapien
   39      436     4.7   1536  4  Q9C0I4                         Q9c0i4 homo sapien
   40    434.5     4.6    454  4  Q96H81                         Q96h81 homo sapien
   41    434.5     4.6   1637  6  Q9XSV8                         Q9xsv8 bos taurus
   42    407.5     4.4   1989  5  Q8I3A0                         Q8i3a0 ciona intes
   43    399.5     4.3    566  5  Q9VKV3                         Q9vkv3 drosophila
   44      378     4.0   4123  4  O75851                         O75851 homo sapien
   45    374.5     4.0    348 11  Q91Z56                          Q91z56 mus musculu


                              ALIGNMENTS


RESULT 1
Q9UL17
ID   Q9UL17      PRELIMINARY;      PRT;   1023 AA.
AC   Q9UL17;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   KIAA1233 protein (Fragment).
GN   KIAA1233.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=BRAIN;
RX   MEDLINE=20039619; PubMed=10574462;
RA   Nagase T., Ishikawa K., Kikuno R., Hirosawa M., Nomura N., Ohara O.;
RT   "Prediction of the coding sequences of unidentified human genes. XV.
RT   The complete sequences of 100 new cDNA clones from brain which code
RT   for large proteins in vitro.";
RL   DNA Res. 6:337-345(1999).
DR   EMBL; AB033059; BAA86547.1; -.
DR   InterPro; IPR003598; Ig_c2.
DR   InterPro; IPR003640; Ig_like.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR000884; TSP1.
DR   Pfam; PF00047; ig; 3.
DR   Pfam; PF00090; tsp_1; 6.
DR   SMART; SM00408; IGc2; 2.
DR   SMART; SM00410; IG_like; 1.
DR   SMART; SM00209; TSP1; 6.
DR   PROSITE; PS50092; TSP1; 1.
KW   Immunoglobulin domain.
FT   NON_TER       1      1
SQ   SEQUENCE   1023 AA;  113521 MW;  D0ABCFF6A0A5D0A1 CRC64;

  Query Match              59.9%;  Score 5599;  DB 4;  Length 1023;
  Best Local Similarity    99.9%;  Pred. No. 0;
  Matches 1022;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        669 AVCLHIQTQQTVNDSLCDMVHRPPAMSQACNTEPCPPRWHVGSWGPCSATCGVGIQTRDV 728
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AVCLHIQTQQTVNDSLCDMVHRPPAMSQACNTEPCPPRWHVGSWGPCSATCGVGIQTRDV 60

Qy        729 YCLHPGETPAPPEECRDEKPHALQACNQFDCPPGWHIEEWQQCSRTCGGGTQNRRVTCRQ 788
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 YCLHPGETPAPPEECRDEKPHALQACNQFDCPPGWHIEEWQQCSRTCGGGTQNRRVTCRQ 120

Qy        789 LLTDGSFLNLSDELCQGPKASSHKSCARTDCPPHLAVGDWSKCSVSCGVGIQRRKQVCQR 848
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 LLTDGSFLNLSDELCQGPKASSHKSCARTDCPPHLAVGDWSKCSVSCGVGIQRRKQVCQR 180

Qy        849 LAAKGRRIPLSEMMCRDLPGFPLVRSCQMPECSKIKSEMKTKLGEQGPQILSVQRVYIQT 908
              |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db        181 LAAKGRRIPLSEMMCRDLPGLPLVRSCQMPECSKIKSEMKTKLGEQGPQILSVQRVYIQT 240

Qy        909 REEKRINLTIGSRAYLLPNTSVIIKCPVRRFQKSLIQWEKDGRCLQNSKRLGIIKSGSLK 968
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 REEKRINLTIGSRAYLLPNTSVIIKCPVRRFQKSLIQWEKDGRCLQNSKRLGITKSGSLK 300

Qy        969 IHGLAAPDIGVYRCIAGSAQETVVLKLIGTDNRLIARPALREPMREYPGMDHSEANSLGV 1028
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 IHGLAAPDIGVYRCIAGSAQETVVLKLIGTDNRLIARPALREPMREYPGMDHSEANSLGV 360

Qy       1029 TWHKMRQMWNNKNDLYLDDDHISNQPFLRALLGHCSNSAGSTNSWELKNKQFEAAVKQGA 1089
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TWHKMRQMWNNKNDLYLDDDHISNQPFLRALLGHCSNSAGSTNSWELKNKQFEAAVKQGA 420

Qy       1089 YSMDTAQFDELIRNMSQLMETGEVSDDLASQLIYQLVAELAKAQPTHMQWRGIQEETPPA 1148
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 YSMDTAQFDELIRNMSQLMETGEVSDDLASQLIYQLVAELAKAQPTHMQWRGIQEETPPA 480

Qy       1149 AQLRGETGSVSQSSHAKNSGKLTFKPKGPVLMRQSQPPSISFNKTINSRIGNTVYITKRT 1208
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AQLRGETGSVSQSSHAKNSGKLTFKPKGPVLMRQSQPPSISFNKTINSRIGNTVYITKRT 540

Qy       1209 EVINILCDLITPSEATYTWTKDGTLLQPSVKIILDGTGKIQIQNPTRKEQGIYECSVANH 1268
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 EVINILCDLITPSEATYTWTKDGTLLQPSVKIILDGTGKIQIQNPTRKEQGIYECSVANH 600

Qy       1269 LGSDVESSSVLYAEAPVILSVERNITKPEHNHLSVVVGGIVEAALGANVTIRCPVKGVPQ 1328
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 LGSDVESSSVLYAEAPVILSVERNITKPEHNHLSVVVGGIVEAALGANVTIRCPVKGVPQ 660

Qy       1329 PNITWLKRGGSLSGNVSLLFNGSLLLQNVSLENEGTYVCIATNALGKAVATSVLHLLERR 1388
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 PNITWLKRGGSLSGNVSLLFNGSLLLQNVSLENEGTYVCIATNALGKAVATSVLHLLERR 720

Qy       1389 WPESRIVFLQGHKKYILQATNTRTNSNDPTGEPPPQEPFWEPGNWSHCSATCGHLGARIQ 1448
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 WPESRIVFLQGHKKYILQATNTRTNSNDPTGEPPPQEPFWEPGNWSHCSATCGHLGARIQ 780

Qy       1449 RPQCVMANGQEVSEALCDHLQKPLAGFEPCNIRDCPARWFTSVWSQCSVSCGEGYHSRQV 1508
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 RPQCVMANGQEVSEALCDHLQKPLAGFEPCNIRDCPARWFTSVWSQCSVSCGEGYHSRQV 840

Qy       1509 TCKRTKANGTVQVVSPRACAPKDRPLGRKPCEGHPCVQWEPGNRCPGRCMGRAVRMQQRH 1568
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TCKRTKANGTVQVVSPRACAPKDRPLGRKPCEGHPCVQWEPGNRCPGRCMGRAVRMQQRH 900

Qy       1569 TACQHNSSDSNCDDRKRPTLRRNCTSGACDVCWHTGPWKPCTAACGRGFQSRKVDCIHTR 1628
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TACQHNSSDSNCDDRKRPTLRRNCTSGACDVCWHTGPWKPCTAACGRGFQSRKVDCIHTR 960

Qy       1629 SCKPVAKRHCVQKKKPISWRHCLGPSCDRDCTDTTHYCMFVKHLNLCSLDRYKQRCCQSC 1688
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 SCKPVAKRHCVQKKKPISWRHCLGPSCDRDCTDTTHYCMFVKHLNLCSLDRYKQRCCQSC 1020

Qy       1689 QEG 1691
              |||
Db       1021 QEG 1023
```

```
RESULT 2
P82987
ID   P82987      PRELIMINARY;      PRT;   766 AA.
AC   P82987;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE   ADAM-TSL3 precursor (Fragment).
GN   ADAMTSL3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Hirohata S., Anand-Apte B., Seldin M., Apte S.;
RT   "Punctin, a member of a new family with similarities to ADAM TS
RT   proteases, is a component of extracellular matrix of skeletal
RT   muscle.";
RL   Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: CONTAINS 6 TSP TYPE 1 DOMAINS.
DR   EMBL; AF237652; AAK15041.1; -.
DR   InterPro; IPR000884; TSP1.
DR   Pfam; PF00090; tsp_1; 6.
DR   SMART; SM00209; TSP1; 7.
DR   PROSITE; PS50092; TSP1; 1.
KW   Repeat; Signal.
FT   SIGNAL         1     26       POTENTIAL.
FT   CHAIN         27    766       ADAM-TSL3.
FT   DOMAIN        79    123       TSP TYPE-1 1.
FT   DOMAIN       422    474       TSP TYPE-1 2.
FT   DOMAIN       482    528       TSP TYPE-1 3.
FT   DOMAIN       568    625       TSP TYPE-1 4.
FT   DOMAIN       648    703       TSP TYPE-1 5.
FT   DOMAIN       707    759       TSP TYPE-1 6.
FT   NON_TER      766    766
SQ   SEQUENCE   766 AA;  85987 MW;  A4261BE87AE91719 CRC64;

  Query Match             46.2%;  Score 1322;  DB 4;  Length 766;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches  765;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MASWTSPWWVLIGMVFMHSPLPQTTAEKSPGAYFLPEFALSPQGSFLEDTTGEQFLTYRY 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MASWTSPWWVLIGMVFMHSPLPQTTAEKSPGAYFLPEFALSPQGSFLEDTTGEQFLTYRY 60

Qy         61 HVQISENIKSPDEDKLGNWLAWGDWSKCSRTCGGGASYSLRRCLTGRNCEGQNIRYKTCSN 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 HVQISENIKSPDEDKLGNWLAWGDWSKCSRTCGGGASYSLRRCLTGRNCEGQNIRYKTCSN 120

Qy        121 HDCPPDAEDFRAQQCSAYNDVQYQGHYYEWLPRYNDPAAPCALKCHAQGQNLVVELAPKV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 HDCPPDAEDFRAQQCSAYNDVQYQGHYYEWLPRYNDPAAPCALKCHAQGQNLVVELAPKV 180

Qy        181 LDGTRCNTDSLDMCISGICQAVGCDRQLGSNAKEDNCGVCACDGSTCRLVRGQSKSHVSP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 LDGTRCNTDSLDMCISGICQAVGCDRQLGSNAKEDNCGVCACDGSTCRLVRGQSKSHVSP 240

Qy        241 EKREENVIAVPLGSRSVRITVKGPAHLFIESKTLQGSKGEHSFNSPGVFVVENTTVEFQR 300
              |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db        241 EKREENVIAVPLGSRSVRITVKGPVHLFIESKTLQGSKGEHSFNSPGVFVVENTTVEFQR 300

Qy        301 GSERQTFKIPGPLMADFIFKTRYTAAKDSVVQFFFYQPISHQWRQTDFFPCTVTCGGGYQ 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GSERQTFKIPGPLMADFIFKTRYTAAKDSVVQFFFYQPISHQWRQTDFFPCTVTCGGGYQ 360

Qy        361 LNSAECVDIRLKRVVPDHYCHYYPENVKPKPKLKECSMDPCPSSDGFKEIMPYDHFQPLP 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LNSAECVDIRLKRVVPDHYCHYYPENVKPKPKLKECSMDPCPSSDGFKEIMPYDHFQPLP 420

Qy        421 RWEHNPWTACSVSCGGGIQRRSFVCVEESMHGEIIQVEEWKCMYAPKPKVMQTCNLFDCP 480
```

```
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 RWEHNPWTACSVSCGGGIQRRSFVCVEESMHGEILQVEEWKCMYAPKPKVMQTCNLFDCP 480

Qy        481 KWIAMEWSQCTVTCGRGLRYRVVLCINHRGEHVGGCNPQLKLHIKEECVIPIPCYKPKEK 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 KWIAMEWSQCTVTCGRGLRYRVVLCINHRGEHVGGCNPQIFLHIKEECVIPIPCYKPKEK 540

Qy        541 SPVEAKLPWLKQAQELEETRIATEEPTFIPEPWSACSTTCGPGVQVREVKCRVLLTFTQT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 SPVEAKLPWLKQAQELEETRIATEEPTFIPEPWSACSTTCGPGVQVREVKCRVLLTFTQT 600

Qy        601 ETELPEEECEGPKLPTERPCLLEACDESPASRELDIPLPEDSETTYDWEYAGFTPCTATC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ETELPEEECEGPKLPTERPCLLEACDESPASRELDIPLPEDSETTYDWEYAGFTPCTATC 660

Qy        661 LGGHQEATAVCLHIQTQQTVNDSLCDMVHRPPAMSQACNTEPCPPRWHVGSWGPCSATCG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 LGGHQEATAVCLHIQTQQTVNDSLCDMVHRPPAMSQACNTEPCPPRWHVGSWGPCSATCG 720

Qy        721 VGTQTRDVYCLHPGETPAPPEECRDEKPHALQACNQFDCPPGWHIE 766
              ||||||||||||||||||||||||||||||||||||||||||||||
Db        721 VGTQTRDVYCLHPGETPAPPEECRDEKPHALQACNQFDCPPGWHIE 766


RESULT 3
Q95LQ2
ID   Q95LQ2       PRELIMINARY;       PRT;   761 AA.
AC   Q95LQ2;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Hypothetical 84.2 kDa protein.
OS   Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Macaca.
OX   NCBI_TaxID=9541;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=TESTIS;
RA   Hashimoto K., Osada N., Hida M., Kusuda J., Tanuma R., Hirai M.,
RA   Terao K., Sugano S.;
RT   "Isolation of novel full length cDNA clones from macaque testis cDNA
RT   libraries.";
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB072736; BAB69705.1; -.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR000884; TSP1.
DR   Pfam; PF00047; ig; 3.
DR   Pfam; PF00090; tsp_1; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   761 AA;  84186 MW;  CEB3F558FCB8AF06 CRC64;

  Query Match             39.9%;  Score 3729;  DB 6;  Length 761;
  Best Local Similarity   94.9%;  Pred. No. 7.1e-276;
  Matches  684;  Conservative   17;  Mismatches   20;  Indels    0;  Gaps    0;

Qy        940 QKSLTQWEKDCRCTLNCKKLDSPRSCSFEIHCTAAPCGCVYRCTACSAQLIVVGKLTIDD 999
              :||||||||||||||||||||||||||| || ||||||||||||||||||||:|||||:|
Db         27 FKSSIQWEKDCRCTLNSKKLKSPRSCSLKIHSLAIPSGCVYRCTACSAQLIVVGKLIIDD 86

Qy       1000 NELIAPPALEEPMKEYPGMGHGEANSGSVTWHKMRQMWNNKNDLYLDDGHISNQFFLKAL 1059
              ||||||||||||||||||||:|||||||||||||||||||||||||||||:||||||||||
Db         87 NELIAPPALEEPMKEYPGMGHNEANSGSVTWHKMRQMWNNKNDLYLDDGYISNQFFLKAL 146

Qy       1060 LGHCSNCACGINGWELKNKQPHAAVKQTAYCMGTALPFELIRNMQLMEIGHVSLGLAQ 1119
              |||||:|||:||||||||||||||||||||||||||||||||||||||||||||||||||
Db        147 LGHCSTCACSINTNCWEFKNKQPRAAVKQTAYCMGTALPTELIRNMQLMEIGHVSLGLAQ 206

Qy       1120 LIYQIVALAKAQTTHMQWNNLQLEDPGAAQLRTLTTDVQQQTNAKNSQKLTTKENKVL 1179
              ||||||||||||||||| ||||| |||||||||||||||:||||||||||||||||||
Db        207 LIYQLVALLAKAQPTHMQWNNLQPPMPPAAQLRTLTTDVQQQTNAKNSQKLTTKEKDVL 266
```

```
Qy       1180 MRQSQPPSTSFNKTINSRPTNIVYTPKRTEVINTTCPLTTPSEATYIWIKPCILLQPSVK 1239
              |||||||||||||||||||||||||||||||||||||||||| |||||||.||||||||||
Db        267 MRQSQPPSTCFNKTINSRPTNIVYTTKRTEVINTTCPLTTPCEATYIWIKPCILLQPSVK 326

Qy       1240 TTLDCTGKTQTQNPTRKEQCIYECSVANHLGSDVLSSSVLYAEAPVTLSVEKNTTKPEHN 1299
              |||||||||||||||||||||||||||||||||||||||||||||||||:|| ||||||
Db        327 TTLDCTGKTQTQNPTRKEQGIYECSVANHLGSDVESSSVLYAEAPVTLSIERNTTKPEHN 386

Qy       1300 HLSVVVGGIVEAALGANVTIRCPVKGVPQPNTTWLKRGGSLSGNVSLLFNGSLLLQNVSL 1359
              ||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        387 HLSVMVGGIVEAALGANVTIRCPVKGVPQPNTTWLKRGGSLSGNVSLLFNGSLLLQNVSL 446

Qy       1360 ENEGTYVCIATNALGKAVATSVLHLLERRWPESRIVFLQGHKKYILQATNTRINSNDPTG 1419
              |||||||| ||||||||||||||||||||||:||| |||||:|||||||||||||||||||
Db        447 ENEGTYVCTATNALGKAVATSVLHLLERRWPESKIVFTQGHKKYVLQATNTRINCNDPTG 506

Qy       1420 EPPPQEPFWEPGNWSHCSATCGHLGARIQRPQCVMANGQEVSEALCDHLQKPLAGFEPCN 1479
              || ||||||||||| ||||||||||||||||||||||||||||||||||||| ||:||||
Db        507 EPLPQEPFWEPGNWSRCSATCGHLGARIQRPQCVMANGQEVSEALCDHLQKPLVGFQPCN 566

Qy       1480 IRDCPARWFTSVWSQCSVSCGEGYHSRQVTCKRTKANGTVQVVSPRACAPKDRPLGRKPC 1539
              |||||||||||:||||||||||||||||| ||:|||||||||||||||||||||||||||||
Db        567 IRDCPARWFTSMWSQCSVSCGEGYHSRQVMCKQTKANGTVQVVSPRACAPKDRPLGRKPC 626

Qy       1540 FGHPCVQWEPGNRCPGRCMGRAVRMQQRHTACQHNGSDSNCDDRKRPTLRRNCTSGACDV 1599
              |||||||||||||:||||||| |||||| |||||| ||||:|||:|| ||||||||||||
Db        627 FGHPCVQWEPGNQCPGRCMGHAVRMQQRHTACQHNGSDSNCDDRKRPALRRNCTSGACDV 686

Qy       1600 CWHTGPWKPCTAACGRGFQSRKVDCIHTRSCKPVAKRHCVQKKKPISWRHCLGPSCDRDC 1659
              ||| ||||||||||||| ||||||||||||||||||:||||:||||:|||||||||||
Db        687 CWHACPWKPCTAACGRGLQSRKVDCIHTRSCKPVAKRHCVQKKKPVSWRHCLGPSCDSTY 746

Qy       1660 T 1660
              |
Db        747 T 747


RESULT 4
Q96RW4
ID   Q96RW4       PRELIMINARY;       PRT;   525 AA.
AC   Q96RW4;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   ADAM-TS related protein 1.
GN   ADAMTSR1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Hirohata S., Apte S.S.;
RT   "A novel member of ADAM-TS related gene, ADAM-TSR1 (A Disintegrin like
RT   And Metalloproteinase domain with ThromboSpondin type I modules
RT   Related gene 1).";
RL   Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF176313; AAK84170.1; -.
DR   InterPro; IPR000884; TSP1.
DR   Pfam; PF00090; tsp_1; 3.
DR   PROSITE; PS50092; TSP1; 1.
SQ   SEQUENCE   525 AA;  58351 MW;  3F1606C5AAA7BA68 CRC64;

  Query Match             21.9%;  Score 2046;  DB 4;  Length 525;
  Best Local Similarity   68.3%;  Pred. No. 1e-147;
  Matches  344;  Conservative   72;  Mismatches   88;  Indels    0;  Gaps    0;

Qy         63 [illegible] CNHDC 123
              :||  ||:||:|| ||||| ||:|||||||||: ||||||: ::|:|:||||:|||| ||
Db         22 [illegible] CNVDC 81

Qy        124 [illegible] NLVVELAPKVLDG 183
              ||:| ||||||||:|||: | :||||| ||| ||:||| |:| ||||||||||
Db         82 PPEACDPRAQQCSAHNDVKHHCQPYEWLPVSNDPDNPCSLKCQAEGIILVVLLAPKVLDG 141

Qy        184 TRCNTDSLDMCISGICQAVGCDRQIGSNAKEDNCGVCAGDGSTCRLVRGQSKSHVSPEKR 243
              ||| |:||||||||:||  ||||  ||||    |||||||  |||||||||||| || :|  |
Db        142 TRCYTPSIIMCISGICQIVGCDHQLGSIVKEDNCGVCNGDGSTCRLVRGQYKSQLSATKS 201

Qy        244 EENVIAVPLGSRSVRIIVKGPAHLFIESKTLQGSKGEHSFNSPGVFVVENTTVEFQRGSE 303
              :: |:|:| ||| :|: :||  ||::|:|||||:||  :| :|  |  |:|:|::|:||:  :
Db        202 DDIVVAIPYGSRHIRLVLKGPDHLYLETKTLQGTKGENSLSSTGTFLVDNSSVDFQKFPD 261

Qy        304 RQTFKIPGPLMADFIFKTRYIAAKDSVVQEEFYQPISHQWRQTDPRPCIVTCGGGYQLNS 363
              ::  :: |:| |||| | |  :  : || ||| ||||| |:||:|||||: ||||||| |
Db        262 KEILRMAGPLTADFIVKIRNSGSADSTVQPIFYQPIIHRWRETDFFPCSATCGGGYQLTS 321

Qy        364 AECVDIRLKRVVPDHYCHYYPENVKPKPKLKECSMLPCPSSDGFKEIMPYDHFQPLPRWE 423
              ||| |:|   ||| | ||||||||:||||||:||::| ||:|||:|:|||||  : ||||||
Db        322 AECYDLRSNKVVADQYCHYYPENIKPKPKLQECNLDPCPASDGYKQIMPYDLYHPLPRWE 381

Qy        424 HNPWTACSVSCGGGIQRRSFVCVEESMHGEILQVEEWKCMYAPKPKVMQTCNLFDCPKWI 483
                ||||||  |||||| |:  |||| : | :  |||||||| ||   : | ||:||||||:
Db        382 ATPWTACSSSCGGGIQSRAVSCVEEDIQGHVTSVEEWKCMYIPKMPIAQPCNIFDCPKWL 441

Qy        484 AMEWSQCTVTCGRGLRYRVVLCINHRGEHVGGCNPQLKLHIKEECVIPIPCYKPKEKSPV 543
              | ||| ||||||:||||||||||:||| | |||:|: | ||||||::| ||||||| ||
Db        442 AQEWSPCTVTCGQGLRYRVVLCIDHRGMHTGGCSPKTKPHIKEECIVPIPCYKPKEKLPV 501

Qy        544 EAKLPWLKQAQELEETRIATEEPT 567
              |||||| |||||||     :|||:
Db        502 EAKLPWFKQAQELEEGAAVSEEPS 525


RESULT 5
Q9BXY3
ID   Q9BXY3       PRELIMINARY;      PRT;   417 AA.
AC   Q9BXY3;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE   Thrombospondin.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Mao Y., Xie Y., Zhou Z., Zhao W., Zhao S., Wang W., Huang Y., Wang S.,
RA   Tang R., Chen X., Wu C.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF251058; AAK34948.1; -.
DR   InterPro; IPR000884; TSP1.
DR   Pfam; PF00090; tsp_1; 3.
DR   SMART; SM00209; TSP1; 5.
SQ   SEQUENCE   417 AA;  46274 MW;  [illegible] CRC64;

  Query Match             15.9%;  Score 1483;  DB 4;  Length 417;
  Best Local Similarity   60.5%;  Pred. No. 1e-104;
  Matches  254;  Conservative   55;  Mismatches   95;  Indels   16;  Gaps    3;

Qy        311 GPLMADFIFKTRYIAAKDSVVQEEFYQPISHQWRQTDPRPCIVTCGGGYQLNSAECVDIR 370
              ||| |||| | |  :  : || ||| ||||| |:||:|||||: ||||||| |||| |:|
Db          3 GPLTADFIVKIRNSGSADSTVQPIFYQPIIHRWRETDFFPCSATCGGGYQLTSAECYDLR 62

Qy        371 LKRVVPDHYCHYYPENVKPKPKLKECSMDPCPSSDGFKEIMPYDHFQPLPRWEHNPWTAC 430
                ||| | ||||||||:||||||:||::||||:|||:|:|||||  : ||||||   |||||
Db         63 SNKVVADQYCHYYPENIKPKPKLQECNLDPCPASDGYKQIMPYDLYHPLPRWEATPWTAC 122

Qy        431 SVSCGGGIQRRSFVCVEESMHGEILQVEEWKCMYAPKPKVMQTCNLFDCPKWIAMEWSQC 490
              | ||||||| |:  |||| : | :  |||||||| ||   : | ||:|||||||:| ||| |
Db        123 SSSCGGGIQSRAVSCVEEDIQGHVTSVEEWKCMYIPKMPIAQPCNIFDCPKWLAQEWSPC 182

Qy        491 TVTCGRGLRYRVVLCINHRGEHVGGCNPQLKLHIKEECVIPIPCYKPKEKSPVEAKLPWL 550
              ||||:||||||||||:||| | |||:|: | ||||||::| ||||||| |||||||
Db        183 TVTCGQGLRYRVVLCIDHRGMHTGGCSPKTKPHIKEECIVPIPCYKPKEKLPVEAKLPWF 242

Qy        551 KQAQELEETRIATEEPTFIPEPWSACSTTCGPGVQVREVKCRVLLTFTQTETELPEEECE 610
              ||||||||     :|||:||||  ||||:  |||  | |||  |:|:|||:|:|:    :|| :|||
Db        243 KQAQELEEGAAVSEEPSFIPEAWSACTVTCGVGTQVRIVRCQVLLSFSQSVADLPIDECE 302

Qy        611 GPKLPTERPCLLEACDESPASRELDIPLP----------EDSETIYDWEYAGFTPCTATC 660
              |||   ::|          ||     | | | |:       |        :|  :   ||||| ||| |:  :|
Db        303 GPKPASQP       ACYAGPCSGEIPFENPDETDGLEGGLQDEDELYDWEYEGFTKCSESC 357

Qy        661 LGGHQEAIAVCLHIQTQQTVNDSLCDMVHRPPAMSQACNTEPCPPRWHVGS-WGPCSATC 719
               || |||:   ||: ||::     ::||       |||  :  ::||  :||| !   :  |  |   |:   |
Db        358 GGGVQEAVVSCLNKQTREPAEENLCVTSRRPPQLLKSCNLDPCPARSSIDSAWNACNVLC 417


RESULT 6
Q95R33
ID   Q95R33       PRELIMINARY;      PRT;  1014 AA.
AC   Q95R33;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   GH19218p.
GN   CG2131 OR CG2132.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RA   Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA   Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA   Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA   Nunoo J., Pacleb J., Paragas V., Park S., Phouanenavong S., Wan K.,
RA   Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL   Submitted (NOV-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY061825; AAL27636.1; -.
DR   FlyBase; FBgn0032959; CG2131.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR000884; TSP1.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00090; tsp_1; 6.
DR   PROSITE; PS50092; TSP1; 2.
SQ   SEQUENCE   1014 AA;  112241 MW;  030A1645935D5360 CRC64;

  Query Match             13.6%;  Score 1270.5;  DB 5;  Length 1014;
  Best Local Similarity   26.0%;  Pred. No. 6.9e-88;
  Matches  297;  Conservative  138;  Mismatches  324;  Indels  383;  Gaps   25;

Qy          3 SWTSPWWVLIGMVFMHSPLPQTTAFKSPGAYFLPFFA-LSPQGSFLEDTTGEQFLTYRYD 61
              :||||                       |    | |:        |  | :: :           |        :|
Db          6 NWTSP----------------------KIWAALFVTTLCLLPTQAAWAKKLNASQ----AFD 41

Qy         62 DQTSRNTRSDEDKD-----------------------GNWDAWGDWSDCSRTCGGGASYSL 99
              |   : || || : :                        | |  :| ||| ||||| ||   : :
Db         42 D  [illegible]WSDWSTCSRTCDGGIMHQM 99

Qy        100 RRT[illegible]AYNDVQTQGHYYEWLERYNDPAA 159
              ||| :   :| |:: ||: |:     | |: :|||: |||||||| | |   |:| | | |
Db        100 RR[illegible]AINDVTSLGTLYKWIPHY-DYVE 157

Qy        160 PCALKC------------------------HAQGQNLVVELAPKVLDGTRCNTDSLDMCI 195
              |||| |                        |   |:::|:|:  :| ||||| : ||||||
Db        158 PCALP[illegible]SLDMCI 217

Qy        196 SGICQAVGCDRQIGSNAKEDNCGVCAGDGSTCRLVRGQSKSHVSPEKREENVIAVPLGSR 255
               | || |||| ::||   | | |||| |||::|
Db        219 QGKCQPVCGDLKIGSTKKIDGGVCGGDGNSCS                            250

Qy        256 SVRIIVKGPAHLFIESKTLQGSKGEHSFNSPGVFVVENTTVEFQRGSERQTFKIPGPLMA 315

Db        251 ------------------------------------------------------------ 250
```

```
Qy        316 DLIFKIFYIAAKDSVVQFFFYQPIGHQWRQIIFTIPCTVIPPGIYQGNGAECVIDIRLKEVV 375

Db        251                           QPL FNWEMAPMSQCSVIPCSSYYKMSRPPPENRLINADV 288

Qy        376 PDHYCHYYFENVKPKPKLKECSMPPCPSSPGFKEIMPYDHFQPLPPWEHNPWIACSVSCG 435

Db        289 DDIICSV    INFPEASVEQCNIHSCP                    PRWIADDWSICSRICG 328

Qy        436 GGIQRRGFVCVPESMHGEIIQVEEWKCMYAPKPKVMQICNLFDCPKWIAMEWSQCTVTCG 495

Db        329 HGYRERMVVCAEESNGIKIPVADIMCP IPKPPIQETCIIEECPHWEVEEWIGCSVSCG 386

Qy        496 RGLRYRVVLCINHRGEHVGGCNPQLKLHIKEECVIPIPC                      534

Db        387 QGIQMRGVECKSIDGSTLSAKCPPLTKPPSMQQCSICIHPCGGSLNKVGCTEIVGSSRGGNE 446

Qy        535                                              YKPKEKSPVEAK 546

Db        447 RGERQLEGSDADEDNEDENDESDDVDDLESGQDTDDGEGLSYADQFLLYAHRIQSRLNQE 506

Qy        547 LPWLKQAQEL               FEIRIAIEEPIEIPE PWSACSIICGPGVQVREVKCR 592

Db        507 APDEPFIMHEMNGNGNNNENEGEDESEGPSLEPIYIKDNEWSPCSVICGEGIRRRIYNCK 566

Qy        593 VLLIFTQTETELEEEECEGPKLPTERPCLLEACDESPA    SREEDTPLEEDS       642

Db        567 IFLEYSRIVAIVNDSLCEGKKPHDE     VERCVEDPCMLPSHGFDDQFPRDSIKVGVSE 622

Qy        643 EIIYDWEYAGFIPCIAICGSGHQEAIAVCLHIQTQQIVNDSICDMVHRPEAMSQACNIE 701

Db        623 PGKIYVWKEQGYIGGSAGCLGVEELIINCVREDNGKVVSPELCSPEIKPEARVKICNGR 682

Qy        702 PCPPRWHVGSWGPCSAICGVTIQTRDVYCLHP     GHIP   APTEECRDEKIHALQACN 755

Db        683 PCPPRWNYSEYIECSKSCGSGIKIREVQCIHEVIRGGDNTMVVENGMCTGPFPADRQYCN 742

Qy        756 QFDCPPGWHIEEWQQCSR[illegible]QNRRV[illegible]NL[illegible]QPPKASSHRC[illegible]A 815

Db        743 V[illegible]PVRWEV[illegible]EWC[illegible]MA[illegible]M[illegible]AREA[illegible]N [illegible]

Qy        816 PIDCPPPLAVCPWCP[illegible]V[illegible]QPRR[illegible]LV[illegible]AA[illegible]LVE[illegible] [illegible]

Db        803 VKPCPPE                                                     809

Qy        876 QMPECSKTKSPMKTKI[illegible]KAYI[illegible]V[illegible]K[illegible] [illegible]

Db        810                     DPKTVTQTNN[illegible]FV[illegible]A[illegible]VE[illegible]QVE[illegible] [illegible]

Qy        935 PVP[illegible]QWEK[illegible]N[illegible]AAS[illegible]A[illegible]AL[illegible]VV [illegible]

Db        853 PVKRYNPIKIKWCK[illegible]KV[illegible]A[illegible]VA[illegible]AP[illegible] [illegible]

Qy        993 LK 994

Db        913 VK 914
```

```
RESULT 7
Q19791
ID   Q19791         PRELIMINARY;      PRT;   2165 AA.
AC   Q19791; Q27524;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   F25H8.3 protein.
GN   F25H8.3.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Gajadsty S.;
RL   Submitted (FEB-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94150718; PubMed=7906398;
RA   Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA   Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA   Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA   Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA   Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA   Lightning J., Lloyd C., Mcmurray A., Mortimore B., O'Callaghan M.,
RA   Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
RA   Smaldon N., Smith A., Sonnhammer E., Staden R., Sulston J.,
RA   Thierry Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA   Watson A., Weinstock L., Wilkinson Sproat J., Wohldman P.;
RT   "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT   elegans.";
RL   Nature 368:32-38(1994).
DR   EMBL; Z69361; CAA93288.1; -.
DR   EMBL; Z69360; CAA93288.1; JOINED.
DR   EMBL; Z69360; CAA93287.1; -.
DR   EMBL; Z69361; CAA93287.1; JOINED.
DR   HSSP; P15167; 1DTH.
DR   MEROPS; M12.135; -.
DR   InterPro; IPR001590; Reprolysin.
DR   InterPro; IPR000884; TSP1.
DR   InterPro; IPR000130; Zn_MIpeptdse.
DR   Pfam; PF01421; Reprolysin; 1.
DR   Pfam; PF00090; tsp_1; 14.
DR   SMART; SM00209; TSP1; 18.
DR   PROSITE; PS50215; ADAM_MEPRO; 1.
DR   PROSITE; PS50092; TSP1; 6.
DR   PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
SQ   SEQUENCE   2165 AA;  244097 MW;  [illegible]AAAA9C4888 CRC64;

  Query Match              12.7%;  Score 1183.5;  DB 5;  Length 2165;
  Best Local Similarity    22.7%;  Pred. No. 9.6e-81;
  Matches  402;  Conservative  212;  Mismatches  [illegible];  Indels  589;  Gaps   80;

Qy         79 DGNWDAWGDW[illegible]GADY[illegible]      GPNC[illegible]NIRYKI[illegible]GNHE[illegible]PPLAE 128

Db        603 DGQWGDWKGWGE[illegible]VQRG[illegible]KEKNG[illegible]KY[illegible]VG[illegible]RLRYRSCNIQECPWDTQ 662

Qy        129 DFRAQQCSAYNK DVQYQG        HYYEWLPRYND  PAAPCALKCHAQGQNLVVELAP 178

Db        663 FYKEVQ[illegible]GEFNNK[illegible]IQ[illegible]VAGINIH    WVPKYANVAPNERCKLYCRLSGSAAFYLLRD 719

Qy        179 KVLD[illegible]N[illegible]LEMC[illegible]QAV[illegible]RQ[illegible]NAKLINC[illegible]V[illegible]A[illegible]RLVP[illegible]KSHV 238

Db        720 KVV[illegible]VA[illegible]MPA[illegible]LRR[illegible]K[illegible]V[illegible]GDSSCKVVKGIENEQG 779

Qy        239 CPEKKEENYIAVE[illegible]VK[illegible]ARLF[illegible]K    [illegible]SKGERSENSPGVEVV        291

Db        780 IPGYNE   VMKIPA[illegible]SANI[illegible]RQKGYNNMKE[illegible]NYL[illegible]RAANGEPLLN   GHEQVSGLARQ 839

Qy        292       ENTIVEFQRGSERQIFKI   PGPLMADF                    IFKIFYIAAKDSVVQ 332

Db        839 QIAFQDIVLEYS GSDAIIEKINGIGPIKSGIYVHVISVKSHPPIGYEYMCAAVENAV  893

Qy        333 FFFYQPIG    HQWRQIIF IP TVI[illegible]NGAE[illegible]VI[illegible]LKEVV[illegible]YYFENV   387

Db        894      IRPISSALYLWPVI[illegible]WIE[illegible]PA[illegible]KLE[illegible]HE[illegible]PN     QNVL 945

Qy        388 KPKPKLKECSMP                                PCPSSPGFKEIMPYDHF       416

Db        946 KPKQATRMCNIDPCSIRWIIES[illegible]AK[illegible]KP[illegible]EVSCVKME[illegible]PQIPASEHLCDRNS 1005

Qy        417 QPLP                PWEHNPWIACSVSCG  GGIQRRGFVCVLESMHGEIIQVEEWKCM 463

Db       1006 KPSDIASCYIDCSGRKWNYGEWI[illegible]ELT[illegible]NGKMHEKSY [illegible]VERGNE    RVDESLCG 1060

Qy        464 YAPKPKVMQICNLFDCPKWIAMEWSQCTVTCGRGLRYRVVLCIN  HRGEHVGGCNP QL 520
```

```
Db   1061 KEQKEATEKEINFIFCLKWVYGHWSECSKSCEGCVKMKHAQCI [illegible]AAEKECIHISK[illegible]A[illegible] 1120

Qy    521 KIHIFKEECVIPIPCYKPKEKSPVEAKLPWLKQAQELEETRIAIEEPTFIPEPWSACSTTC 580
           : |  |                              | :                             || ||   |
Db   1121 QEHCNEHA------                  CTWWQFG------------------VWSDCSAKC 1144

Qy    581 GPGVQVREVKCRVLLTFTQTETELPEEECEGPKLPTERPCLLE---------------AC 625
           | ||| |:   |           :   : |||  |    :      :||  |                :|
Db   1145 GDGVQYRDANCI----  DRHRSVIPEHRCLKMEKIIIKPCHRESCPKYKLGEWSQCSVSC 1199

Qy    626 DESPASR-----------ELDIPL-------PEDSETTY-----DWEYAGFTPCIAICLGG 663
           ::   :||            |:|: |          |    :|          |     :: |: :|   |
Db   1200 EDGWSSRKVSCVSGNGIEVDMSLCGTASDRPASHQTCNLGICPFWRNTDWSACSVSCGIC 1259

Qy    664 HQEAIAVCLHIQTQQTVNDSICDMVHRPPAMSQACNIEPCPPRWHVGSWGPCSATCGVGI 723
           |:|      |::     :|:|:: | |       : |   || |:    ||     |     | ||| ||| ||
Db   1260 HREKTTECIY--KEQSVDASFCGDT-KMPETSQICHLLPC-ISWKPSHWSPCSVICGSGI 1315

Qy    724 QTRDVYCLHPTPTTAPPKE T EDEKPHALQACN   -          QELCPP-GWHIEE 767
           ||| | |     |       |   |  | |: :|     : |                   |  | ||    |
Db   1316 QIRSVSCTRGSFGTIVDEYECDPNTRPRLKKTCEKPTCDGPRVIQKIQADVPPIRWATGP 1375

Qy    768 WQQCSPTCGGGTQNRPVTCPQLLTDGSFLNLSDELCQG-PKASSHKSCARTDCPPHLAVG 826
           |   || ||| ||| |  : ||    : |          | || |        |    |  ::|     ||    :   :
Db   1376 WIACSAICGNGIQRRIIKCRDHVPD-----IPDEYCNHIDKEVSTRNCPLRDC-SYWKMA 1429

Qy    827 DWSKCSVSCGVGIQRRKQVCQRLAAKGRRIPLSEMMCRDLPGFPL-VRSCQMPECSKIKS 885
           :|  :|    :||    :|:  :  |       |     |   |     |  : :   |  |:      |        |:|::    | |
Db   1430 EWEECPATCGTHVQQSRNVTCVSAEDGGRTILKDVIC DVQKRPISARNCRLEPCPK     1485

Qy    886 EMKTPLGEL[illegible]LKEVLRNSKQ[illegible]KLLEKIRLESPAE[illegible]N[illegible]VEEQKS[illegible]   [illegible]
                   ||:                                             |||   ::: :  |     ||
Db   1486 -----GEE--------------------------HIGS--WLIGDWS---KC----------- 1501

Qy    946 WEKPGPCLQNSKPIGIIKSDSLKIHGLAAPDICVYPCIACSAQETVVLKLIGTDNRLIAR 1005
                |        :|   ::                               | :  |    ||      |:       |    :
Db   1502    SASCGGGWRRRSVS  ------------------CISSSCDEIRKKPKMPDKCNEELCP 1539

Qy   1006 PALREPMREYPGMDHSEANSLCVTWHKMPQMWNNKNDLYLDDEHISNQPPLRALLCHCSN 1065
           |         :   |                                    |                             |||
Db   1540 PLINNSWQISP------ ----------------WT--------------------------HCSV 1556

Qy   1066 SAGSTNSWELKNKQPFAAVKQGAYSMDTAQFDELIRNMSQLMETGEVSDDLASQLIYQLV 1125
           |  |                          |             ::::          :|     ||:
Db   1557 SCG  ---------------GGVQPRKIWCEDVL---------SCRKQDDI---------- 1581

Qy   1126 AELAKAQPTHMQWRGIQEEIPPA AQIRGPTGSVSQSSHAKNSCKLTFKPKG    PVLME 1181
            | :: :|     :           | ||    :         :| | | :| : ::       |             |:|
Db   1582 -ECSEIKPREQR ---DCEMPPCRSHYHNKTSSASMTSLSSSNSNTTSSASASSLPIL-- 1634

Qy   1182 QSQPPSISPN     KTINSRIGNTVYITKE  TEVINILCDLITPSEATYTWTKDGTLLQ 1235
                 || :|:            ::: |         |||     | :|
Db   1635 ---PPVVSWQTSAWSACSAKCGRG---TKRRVVECVN------------------------ 1665

Qy   1236 PSVKIIIDGTGKIQIQNPTKKEQGIYKCSVANHLG[illegible]VESSVLVAFAPVILSVTPNITK 1295
           ||:  :  :   |      |  :  |    :
Db   1666 PSLNVIVASIECDQIKKPVPP         ------       -----  ----      --  1686

Qy   1296 PEHNHLSVVVGGIVEAALGANVTIRCPVKGVPQ-PNITWLKRGGSLSGNVSLLFNGSLLL 1354
                                    :||   |   |:      ||
Db   1687 ------------------------VRCRTKHCPRWKTTTWSS---------------------- 1704

Qy   1355 QNVCLENEGTYVCIATNALGKAVATSVLHLLEPEWPESKIVFLQGHEKYILQA TNIFTN 1413
                      |    |      |                   : ||        :      :| |   :    :    |  :|
Db   1705                 CSVTCGRG     --         IRRR     EVQCYRGRKNLVSDSECNPKIK 1738

Qy   1414 SNDPIGFPPPQEPE--WEPGNWSHCSATCGHLGARIQRPQCVMANGQEVSEALCDHLQKP 1473
            |           |       |       |       || |      |            : :|    |:       :|:       :    :|:     :  |
Db   1739 LNGVANCEEVACPATEWNVTEWSKCKLECARGQKQTRRVHCTSTSGKPAAPPMCELAPAP 1798

Qy   1474 LAGFEPCNIRPCPARWFTSVWSQCSVSCGEGYHSRQVTCKR TKANGTVQVV   SPRAC 1527
                 |:    :||       |     ||    |||||    :|:.   |:|             |  |:   ::
Db   1799 -ISIRECHPCNTPYEWVIGEWQTCSKSCGKGVQTRRVRCHRKLNENSIIPIIEMLEDEPA 1857
```

```
Qy   1528 APKDRDLGRKPCPGHPCVQWETGNRGPGRCMGRAVKMQQRHTACQHNGSDSNCDDRKRPT 1587
            ||::                                                                    |:     :|
Db   1858 VPKEK--------------------------------------------CELPPKPN 1870

Qy   1588 IRRNCTSGACD--VCWHTGPWKPCTAACGRGFQSRKVDCIHTRSCKPVAKRHCVQKKKPI 1645
                :  |        ||         |       |||     |:     ||:|  :  |:| |:             : |  :     |       |||
Db   1871 ESQTCELNPCDSEKEWS[illegible]CPWC[illegible]SKN[illegible]KRR[illegible]VK[illegible]V-ANE[illegible]V[illegible]RVK[illegible]-IKKPR 1928

Qy   1646 SWRHCLGPSCDRDCIDTTHYCMFVKHLNLCSLD 1678
            ::|          :|:|    :|    |    :|    |:  :   |
Db   1929 RTQYCP-    PRNCLPST  CQPLKSQNVKAKD 1955
```

```
RESULT 8
P90884
ID   P90884       PRELIMINARY;      PRT:  1059 AA.
AC   P90884;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-JAN-1999 (TrEMBLrel. 09, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   F53B6.2 protein.
GN   F53B6.2.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   White S.;
RL   Submitted (JUL-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99069613; PubMed=9851916;
RA   none;
RT   "Genome sequence of the nematode C.elegans: A platform for
RT   investigating biology.";
RL   Science 282:2012-2018(1998).
DR   EMBL; Z81086; CAB03121.1; -.
DR   InterPro; IPR003598; Ig_c2.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR000884; TSP1.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00090; tsp_1; 1.
DR   SMART; SM00408; IGc2; 1.
DR   SMART; SM00209; TSP1; 8.
DR   PROSITE; PS50092; TSP1; 2.
KW   Immunoglobulin domain.
SQ   SEQUENCE   1059 AA;  117768 MW;  5B182A42E41C8597 CRC64;

  Query Match             11.9%;  Score 1110.5;  DB 5;  Length 1059;
  Best Local Similarity   21.1%;  Pred. No. 1.2e-75;
  Matches  351;  Conservative  170;  Mismatches  451;  Indels  695;  Gaps   52;

Qy     77 NWLAWGEWSQCS[illegible]CGGGA-Y[illegible]-R[illegible]C[illegible]R[illegible]N[illegible]Q[illegible]NIKYE[illegible]SN[illegible]AP[illegible]AQCS 136
           :| ||    || |::|||||| |     |||||| :  | |:::|:| |:        |       :       |       |
Db     25 SWAAWSIWSE[illegible]IR[illegible]GGGVRE[illegible]RR[illegible]L[illegible]R[illegible]GS[illegible]LSVRI[illegible]V[illegible]A[illegible]N[illegible]R[illegible]SRLA[illegible]G 83

Qy    137 AYNDVQYQGHYYEWLPRYNDPAAPCALKCHAQ--GQNLVVELAPKVLDGTRCN-TDSLDM 193
               ::   :|                              | :  |  ::    |  |  :              :|  |||  |           |    :
Db     84 G-EEIVSRGQ                       CEVVCRSRLTGANFLW     RVDDGTPCQAATGRAV 124

Qy    194 [illegible]QAV[illegible]GGSNAF[illegible]N[illegible]GV[illegible]ELV[illegible]HV[illegible]LENVIAV[illegible] 253
           |   |  ||  ||||    :  |:  :  |  ||||  |   |   ||
Db    125 CSPGSCQIVGCDCLISSSERRDACGVCGGRGDTC-        -----       ----      --- 158

Qy    254 SRSVRITVECPARKITSK[illegible]NSIQVEVVVENT[illegible]VI[illegible]KIPRGPL 313
                                                                   :|
Db    159 -----------------------  -----  -----------DN----------------  ---- 160

Qy    314 MA[illegible]KL  EE[illegible]AAK   [illegible]SVVQ[illegible]WEQ[illegible]Y[illegible]SAE--- 365
                 ||:|       |||        | :|                    |                        | |    :  |
Db    161  GKTIWKVSEEYTACASNCDDIV           DW             SGACRSIAGTSQPT 198
```

```
Qy        364 CVPIPLEPVVEGKYCHYYPLNVEPKPELKDTMLPTSGRGTKEIMPYHHEQPLPAWE 423
                  ||:       ||||:   |       : ::||  : :  |  |    |||                      :|
Db        199 VVCVNAIGGRVVPEKD      ADKLRPKVEARPCPMLTCPS                   EWM 238

Qy        424 HNPWTACSVSCGGGIQRRSFVCVEESMHGEILQVEEWKCMYAPKPKVMQTCNLFDCPKWI 483
                  || |    || |  ::|     ||  : :  |       :  |  :   |      :|      :  |      | :
Db        239 AAEWTECVPSCGEGTKRREVYCV QIAHNVIVHVPDTECLNGTRPAALENCVSISCGRWE 297

Qy        484 AMEWSQCTVTCGRGLRYRVVLCINHRGEHVGG    CNPQLKLHIKEECVIPIPCYKPKEKS 541
              |  :||:||  :||:|:| | | | |           |||  |:    :     :    |    |||         |
Db        298 AGKWSKCTASCGQGVRRRHVAC            VCCSDCDEGCRPRQEITCYAGIPC----S 344

Qy        542 PVEAKLPWLKQAQELEEIRIAT            EEPTFIPEPWSACSTTCGPGVQVREVKCRVL 594
                     | |   :|   |:         :             : |    :      || ||:||| ||    | |:|    :
Db        345 TATNSLDWNGRAY LSGNTPGSMDNHNDWQAPKLVAGEWSTCSSTCGTGVMSRTVECVAV 403

Qy        595 LTFTQTETELPEEECEGPKLPTERPCLLEACDESPASRELDIPLPEDSETTYDWEYAGFT 654
                  :       :||    ||:           |:| |  |:|:           :  |    |  |:  |    |  |  |   :|
Db        404 NPTSSAPTKLPMSECQ     DQEQPKLPESCEVRSCPLQEESKLSED EAPYQWRYGLWT 458

Qy        655 PCTAPCLGGHQEATAVCLHTQTQQTVNDSLCDMVHRPPAMSQACNTEPCPPRWHVGSWGP 714
                |:|:||||  |:|    |:  :  ,  ::|    :  ||     |||    |:  ||    ||||  |           :
Db        459 QCSASCLGGEQKAAIKCTQVSTGKSVQWSQCDAPRRPPEKSRPCNQHPCPPEWLTSKYGS 518

Qy        715 CSATCGVGTQTFDVYCLH                PGETPAPSEECRDEKPHALQACNQFDCPPGWHIE 766
                ||  :||  |        |  |  |                         :  |      :||      :  ||      ||    |
Db        519 CSMSCGSGTAPPSVKCAQIVSKIPGAPAHIVLRDDRCHEKKPQEIEICNVVACPAIW     575

Qy        767 EWQQCSRTCGGGTQNRRVTCRQLLTDGSFLNLSDELCQGPKASSHKSCARTDCPPHLAVG 826
                                :  |:|                                                |      :|                |
Db        576                       VSSTNKEHN                                              ETEKNE       ----LETA 594

Qy        827 DWSKCSVSCGVGIQRRKQVCQRLAAKGRRIPLSEMMCRDLPGFPLVRSCQMPECSKIKSE 886
                 |::||  ||     |  :||:     |:          ::|:            ::  |                        |    |          ||:
Db        595 QWTECSRSCDSGFRRRQVWCETRDSRGKTQRRPDVECDANTKPQTVEVCSPGSCSR---- 650

Qy        887 MPTPTTPQGPQTLVQPEVYTQTSTTPEPINLTTGSPAYLTPNTSVTTP PVSPEQRSTLSW 946
                               |::||    ||:  |   |:|::  |  ||     |  |        ||:  ||||  ::|   |    |  |
Db        651 --              PELLS NRVFEQNALQKKTTTGGGVATTYQTISTKTECPAKKPEKKTYW 700

Qy        947 EKPCRCLQNSKRLGTTKSGSLKIHCLAAPDTCVYRCTACSAQETVVLKLIGTDNRLIARP 1006
                :|:|:  ::|         :  : :  :|:|::                |  |||  |                |    |   |
Db        701 EENGPPTENDAHTEVSANGNLPVPSAKMEDACVYPCTTDRLQGNVTLNP---------- 749

Qy        1007 ALPEPMPEYTPMPHSSEANSLSVTWHKMRQMWNNKNDLYLLEDSHTCNQPPLRALLGHTCNG 1066
                        |::|
Db        750        KYRDPP                                                    -------     - 755

Qy        1067 AGSINSWELKNKQTEAAVKQNAYSMDTAQTDELTKNMSQLMETGEVSDDLASQLLYQLVA 1126

Db        756                                                                    - --- - 755

Qy        1127 ELAKAQPTHMQWRGTQEETTAAQLRGETGSVSQSSHAKNSGKLTFKPKGPVLMRQSQPP 1186
                                                                                  :  ::       ||                      :
Db        756                                                                  ASRVDLAPK                P 765

Qy        1187 STSFNKTTNSRIGNIVYTTKPTEVTNTLCPLTTPSEATYTWTKPTTLLQPSVKTTLDSTG 1246
                   |                                                        ||              ||:         :|  |  :  :|
Db        766 QT                                                        PS         TENPQPVQVSKEEVLR      785

Qy        1247 KTQTQNPTPKEQTTYETSVANHDGSDVESSSVLYAEAPVTLSVERNTTKPEHNHLSVVVG 1306
                                                    |  :||||:                  :  |::
Db        786                                                    EQASVLH                KMNVS                 - 797

Qy        1307 GTVEAALSANVTTRCPVKGVPQPNTTWLKRGGSLSGNVGLLPNGSLLLQNVSLENEGTYV 1366
                    ::|:   |  |                           ||           :            ||
Db        798      LIEALLIA                           PNDEKAREQLRKYGN         -          820

Qy        1367 CTATNALGKAVATSVLHLLRRPWPESTPTVPLQSHKKYTLQATNTRTNCNDPTGRPPPQEP 1426
                                              |:  |
Db        821                              ELVAR                                              825

Qy        1427 FWEPGNWSHCSATCGHLGARTQRPQ  VMANGQ  PVSEALCDHLQ  KPLAGTEPCNT 1480
                  |:  |:||  |      ||      |  :  :      |  |    :|:      |     ::|:  |      :|          ||:
Db        826  WDTGHWSECRQKTCHVACYQARGTSCKVIEHGETRNVDNSTCESLASVKP PETRPCHR 883

Qy        1481 RDCPARWTTGVWGQCSVSCGTGYHGRQVTCKETKANGTVQVVSPPACAPKDRPLGRKPCF 1540
                  |||  ||    |  ||:||
Db        884  EDCP RWEASQWSECSSQ                                                                  900

Qy        1541 GHPCVQWEPGNISTPCSCMCRAVPMQQRRIACHNGSGS          NCSSPKRPTLRRNCTSGA 1596
                                            ||:       :      ::|:     |:     :      |                :|:      ||          :|  :
Db        901                               RCVSSMLAQKRRNVTCETTNGTGVTGQHCTTNRPATTMGTNQN 945

Qy        1597 CDVCWHTGPWEPCTAACGR CEQSPKVDCTHTPSCKPVAKPRCVQKKKPTSWRHCLG      1652
                  |        |   |      |      |:: ||      |  |   |  :  |:        |  :|   |   |:|   |   ::|   |   |   |:
Db        946  CKAEWRTSDWGSCSSECGTGGVQLRLLSCVWTSSGKP AGRNCEQMRKPHSARACVADEP 1004

Qy        1653  PSC                  - DRDCTDTTHYCMFVKHLNLCSLDRYKQRCCQSC 1688
                  | |                        |   |   |  :  :|    :|    :  |              :|:||  :|
Db        1005 LPPCMPTASALYQRDASCQDQSRPCDTTKLFHSCDSLEVRQKCCSTC 1051
```

```
RESULT 9
Q9EPX2
ID   Q9EPX2       PRELIMINARY;       PRT;   1280 AA.
AC   Q9EPX2;
DT   01 MAR 2001 (TrEMBLrel. 16, Created)
DT   01 MAR 2001 (TrEMBLrel. 16, Last sequence update)
DT   01 JUN 2002 (TrEMBLrel. 21, Last annotation update)
DE   Papilin.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID 10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN CD-1;
RX   MEDLINE 20530499; PubMed 11076767;
RA   Kramerova I.A., Kawaguchi N., Nelson R.E., Fessler L.I., Chen Y.,
RA   Kramerov A.A., Kusche-Gullberg M., Kramer J.M., Ackley B.D.,
RA   Sieron A.L., Prockop D.J., Fessler J.H.;
RT   "Papilin in development; a pericellular protein with a homology to the
RT   ADAMTS metalloproteinases.";
RL   Development 127:5475-5485(2000).
DR   EMBL; AF314171; AAG41986.1; -.
DR   HSSP; P12111; 2KNT.
DR   InterPro; IPR000183; Decarbxylse2.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR003598; Ig_c2.
DR   InterPro; IPR003600; Ig like.
DR   InterPro; IPR003006; Ig MHC.
DR   InterPro; IPR002223; Kunitz_BPTI.
DR   InterPro; IPR000884; TSP1.
DR   Pfam; PF00047; ig; 3.
DR   Pfam; PF00014; Kunitz_BPTI; 1.
DR   Pfam; PF00090; tsp_1; 5.
DR   PRINTS; PR00759; BASICPTASE.
DR   ProDom; PD000222; Kunitz BPTI; 1.
DR   SMART; SM00409; IG; 3.
DR   SMART; SM00408; IGc2; 3.
DR   SMART; SM00410; IG_like; 2.
DR   SMART; SM00131; KU; 1.
DR   SMART; SM00209; TSP1; 5.
DR   PROSITE; PS00280; BPTI_KUNITZ_1; 1.
DR   PROSITE; PS50279; BPTI_KUNITZ_2; 1.
DR   PROSITE; PS00879; ODR_DC_2_2; UNKNOWN_1.
DR   PROSITE; PS50092; TSP1; 3.
KW   Immunoglobulin domain; Serine protease inhibitor.
SQ   SEQUENCE   1280 AA;  138824 MW;  AE287705E561AF30 CRC64;

  Query Match             10.3%;  Score 958.5;  DB 11;  Length 1280;
  Best Local Similarity   22.9%;  Pred. No. 6.8e-64;
  Matches  362;  Conservative  182;  Mismatches  572;  Indels  467;  Gaps   66;
```

```
Qy     65 SRNIRSDEDKPCNWDAWCDWSDCSRTCGGGASYSLRRCLT-----GRNCEGQNIRYKTCS 119
          :|| |   |    |  ||:|| |||||||| |:  | | :     | :| |    ::||
Db     20 ARNVRRQSD---TWGTWGEWSPCSRTCGGGISFRERPCYSQRRDGGTSCVGPARSHRTCH 76

Qy    120 NHDCPPDAEDFRAQQCSAYNDVQYQGHYYEWLPRYNDPAAP CALKCHAQGQNLVVELA 177
             ||    ||||:||: ::   :||  | ||| |   |||  | | |  :|||   :
Db     77 IESCPDGVRDFRAEQCAEEDGTDFQGRRYRWLPYY   AAPNKCELNCIPKGQNFYYKHK 133

Qy    178 PKVLPGTPGNTLSLLMCLSGLCQAVGCDRQLGSNAKFLNCGVCACPGSTCRLVPGQSKGH 237
            |:||| |     |::: |:|: ||||:|| :|| |  | ||||:| | | |   :
Db    134 DAVVDGTPCEPGQRDICVDGVCRVVGCDHKLDSIKQEDKCLQCGGDGSSCYPVTCTFDGN 193

Qy    238 VSPEKREENVIAVPLGSRSVRITVKGPAHLFIESKTLQGS   KGEHSFNSPGVFVVENT 294
              :    :  :| |: |:||     :  |:  |::|      |  :  :     |:|
Db    194 -DLSRGYNQIFIIPACATSIRIEEAAASRNFLAVKSIRGEYYLNGHWTIEAAQALPVAST 252

Qy    295 TVEFQRGSF----RQTFKIPGPLMADFIFKTRYTAAKDSVVQFFFYQPIS-----HQWRQ 345
           ::::|| |     :  :  ||     : :   :   :  | : :| | :       |
Db    253 VLQYERGVESDLAFERLQARGPTSEFLVIEL LSQESNFGVHYEYYLPANDPGKGFSWSH 311

Qy    346 TDFFPCIVTCGGGYQLNSAEC-VDIRLKFVVPDHYCHYYPENVKPKPKLKECSMDPCPSS 404
            :  |:  ||||:|     | :|       ||| | : |   :|  : : |:  ||| :
Db    312 GSWGDCSALCGGGHQSRLVPCTID---NEAYPDHMCQHQP   RPTHR RSCNTQPCPKT 364

Qy    405 DGFKEIMPYDHFQPLPRWEHNPWTACSVSCGGGIQRRSFVCVEESMHGEILQVEEWKCM- 463
                          ||:  ||| |||||||:|  ||  |:     |     || :|
Db    365 ---------------KRWKVGPWTPCSVSCGGGVQSRSVYCISSDGTGGQEAAEETQCAG 409

Qy    464 YAPKPKVMQTCNLFDCPKWIAMEWSQCTVTCGRGLRYRVVLCINHRGEHVGGCNPQLKLH 523
           | ||   | |||  |  |    | :|:|||| |:| | | |   ||:
Db    410 LAGKPPTIQACNLQHCAVWSVEPWGECSVTCGTGIRKRSVTC---RGD------------ 454

Qy    524 IKEECVIPIPCYKPKEKSPVEAKLPWLKQAQELEETRIATEEPTFIPE-----PWSACST 578
                         | ||| |    ||    | |  :    | |  :      || ||
Db    455 ---------------EGSPVHAAACLLKPQPTLTEPCVQEACPVTRCQAWHVCNWSLCSK 499

Qy    579 TCGPGVQVRKVKCRVLLIFIQTETELPEEECEGP      KLPTERPCLLEACDESPASRE 633
          :|| |:: |:| | :                 ||      | : :|  :|||:  |
Db    500 SCGSGIRRRQVVCTI-                 GPPGRCVDLQSSKPAEMEACNRQPCHLP 542

Qy    634 LDIPLPEDSETTYDWEYAGFTPCTATCLGGHQEAIAVCLHIQTQQTVNDSLCDMVHRPPA 693
           ::|  :|  |          |  :    | ::  |     : |    :       : |
Db    543 QEVPSIQDPRTRSSPPRILSCPRVSPVSPCRPQQWAPLRPPRAQSNPRRC------QDPN 596

Qy    694 MSQA--CNTEPCPPRWHVGSWGPCSATCGVGIQTRDVYCLHPGETPAPPEECRDEKPHAL 751
          :| |    |   ||  |     | |                       | :||    |:
Db    597 ISSACRAPTIQRPP--HQPPLRPSSQ----------------------PRPCR----HSP 629

Qy    752 QACNQHDCPPGWHL      PEWQQCSRTCGGGTQNRRVTCRQLLTDGSHLNLSDELCQGPK 807
            |    || | |      :|| |        |:|   |     ||          :||:
Db    629 HGC----CPDG-HTPSLGPQWQGCPLAGASCLQSRYGCG----PDGV------SAAFCPQ 673

Qy    808 AS----SHKSCARTDCPPHLAVGDWSKCSVSCGVGIQRRKQVCQRLAAKGRRIPLSEMMC 863
           :    || |    : |   ||               :  :: |  | :|     ||
Db    674 QACCTRSHGSDNIGNRPGSRAVA              SKNPKIHQPQAHEGEP---SE--- 714

Qy    864 RDLPGFPLVRSCQMPECSKIKSEMKTKLGEQGPQILSVQRVYIQTRTEKRINLTIGSRAY 923
                   || :   .    :     |||                        :|  :|
Db    715 ---------CRSSRFGCCYDNVASAAGPLGEG                CVCQPSY 744

Qy    924 LLPNTSVIIKCPVRRFQKSLIQWEKDGRCLQNSKRLGITKSGSLKIHGLAAPDIGVYRCI 983
            |     ::| :   | |   |
Db    745 AYP----VRCLLPSAQGSCGDW------------------------------------- 762

Qy    984 AGSAQETVVLKLIGTDNRLIAPALREEMREYPGMLHSEANSLGVIWHEMSQMWNNKNDDL 1043
             |     :  :|  ||             : |  |  ||:
Db    763 ---AARWYFVASVGRCNRF                WYGGCHGNANNFA--         791

Qy   1044 YLDDPHISNQPFLRALLGHCSNSAGSTNSWELKNKQFEAAVKQGAYSMDTAQFDELIRNM 1103
                 | |         | |:    :      :: ||        :|  |     :
Db    792 -------SFQ---------FCMNICRGQHG----PRRPEACAACHRAHVDGGQRGPGGQQF 832

Qy   1104 SQLMETGEVSDDLASQLIYQLVAPLAKAQPTHMQWRGIQEETPPAAQLRGETGSVSQSSH 1163
```

```
                |     | |              |:   ||  ||  |      ||   :   ::
Db    833 PDWHRAGATIPRLPS--------------PSGSPWRREQEPAP      GEPPHI--PAY 870

Qy   1164 AKNSGKLTFKPKGPVLMRQSQ    FPSISFNKTINGRIGNTVYITKRT   EVINILCDLI 1218
              |    :|: | | |:::   ||: | :  |    |:   : :    : : : |
Db    871 GNRPGGQEIRPRVPGLDREARPAVPPIHSPSYRIR-LAGSEPSLVQAAPGQAVQLPCPGN 929

Qy   1219 IPSEATTWTKLGTLLQPSVKITLSGTGKLQIQNFIRKEQGIYECSV   ANHLGSDVE 1274
           |||    | :|  :  | :  |   | : |     :: |||,      |   :::
Db    930 IPSEFQAGWQKEGRPIS-SNRYQLQADGSLIISRLRPEDAGIYSCGSHRPGHEPQEIQLR 988

Qy   1275 ----SSSVLYAEAPVILSVERN--------------------IIKPEHN------HLSVV 1304
                :||    |      ||                    :  | |        |
Db    989 VTGGDMAVLPEGQPRHEPEPRNPDLGHGPPHRGTGAFAGGHRVLSPSHPRPATRLRLDRT 1048

Qy   1305 VGGIVEAALGANVTIRCPVKGVPQPNITWLKRGGSL   SGNVSLLFNGSLLLQNVSLENE 1362
            |:|:|: |  : :  |  :| | | | | :| | |  |    :  :|||::  | :|:
Db   1049 QPGVVDASPGQRIRLTCRAEGEPVPTIEW-QRDGQLVSSPRHQVQPDGSLVISRVDVEDG 1107

Qy   1363 GTYVCIATNALGKAVATSVLHLLERRWPESRIV---            FLQGHKKYIL---- 1405
          | | |:| |   :          ::| : |::                 :|   :|
Db   1108 GYYSCVAFNGQDR-----------DQRWVQLRVLRELTITGLPPAVIVAEGDTARLLCVVA 1157

Qy   1406 -QATNIRINSNDPTGEEPPQEEEWEEGNWGHCGATCGHLGARTQKPQCVMANGQEVSEA 1463
           :: | | : |   | |                               : |:|  | ::
Db   1158 GESVNIRWSRN---GLP--------------------------   IQADGHRVHQSP 1183

Qy   1464 ---LCDHLQKPLAGFEPCNIRDCPARWFTSVWSQCSVSCGEGYHSRQVTCKRTKANGTVQ 1520
             |  |  :|         ||           : | :|   :   |   : |:    |:
Db   1184 DGTLLIHNLRP----------RD------------EGSYTC-SAERGSQAVSRSTE----VK 1218

Qy   1521 VVSPRACAPKDRPLGRKPCFGHP 1543
          |  | | | : | || | |   |
Db   1219 VALP-APAAQSRDLG-KDCIDQP 1239


RESULT 10
O95428
ID   O95428      PRELIMINARY;      PRT;  1235 AA.
AC   O95428;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Hypothetical 133.5 kDa protein.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Rowen L., Madan A., Qin S., Abbasi N., Dors M., Ratcliffe A.,
RA   Madan A., Dickhoff R., Shaffer T., James R., Lasky S., Hood L.;
RT   "Complete sequence of the gene for presenilin 1.";
RL   Submitted (NOV-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF109907; AAC97963.1; -.
DR   HSSP; P12111; 2KNT.
DR   InterPro; IPR003598; Ig_c2.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR002223; Kunitz_BPTI.
DR   InterPro; IPR000884; TSP1.
DR   Pfam; PF00047; ig; 3.
DR   Pfam; PF00014; Kunitz_BPTI; 1.
DR   Pfam; PF00090; tsp_1; 5.
DR   PRINTS; PR00759; BASICPTASE.
DR   ProDom; PD000222; Kunitz_BPTI; 1.
DR   SMART; SM00408; IGc2; 3.
DR   SMART; SM00131; KU; 1.
DR   SMART; SM00209; TSP1; 5.
DR   PROSITE; PS00280; BPTI_KUNITZ_1; 1.
DR   PROSITE; PS50279; BPTI_KUNITZ_2; 2.
DR   PROSITE; PS50092; TSP1; 4.
KW   Hypothetical protein; Immunoglobulin domain;
KW   Serine protease inhibitor.
```

```
SQ   SEQUENCE   1235 AA;  133477 MW;  A0B440CE4F30E350 CRC64;

  Query Match              9.9%;  Score 913.5;  DB 4;  Length 1235;
  Best Local Similarity   22.7%;  Pred. No. 1.8e-60;
  Matches  350;  Conservative  179;  Mismatches  544;  Indels  471;  Gaps   66;

Qy         78 WDAWGDWSDCSRTCGGGASYSLRRCLF      GRNCEGQNIRYKTCSNHDCPPDAEDFRA 132
              |  |  || ||||||||  |:   | | :     | :| |    :::|    ||  | ||||
Db         29 WGPWSQWSPCSRTCGGGVSFRERPCYSQRRDGGSSCVGPARSHRSCRTESCPDGARDFRA 88

Qy        133 QQCSAYNDVQYQGHYYEWLPRYNDPAAPCALKCHAQGQNLVVELAPKVLDGTRCNTDSLD 192
              :||: ::   ::||  | ||  |: |    | | |  :|:    :     |:|| |     |
Db         89 EQCAEFDGAEFQGRRYRWLPYYSAP NKCELNCIPKGENFYYKHRLAVVDGTPCEPGKRD 147

Qy        193 MCISGICQAVGCDRQLGSNAKEDNCGVCAGDGSTCRLVRGQSKSHVSPEKREENVIAVPL 252
              :|: |  |: ||||  :|  |:  :||  |  | |||:|  ||  ||    ::    :   :: |::
Db        148 VCVDGSCRVVGCDHELDSSKQEDKCLRCGGDGTTCYPVAGTFDAN DLSRGYNQILIVPM 206

Qy        253 GSRSVRITVKGPAHLFIESKTLQGS     KGEHSFNSPGVFVVENTTVEFQRGSE----RQ 305
              |: |: |       :   |:  : ::|     |  :   :           :| : ::|:|:|     :
Db        207 GATSILIDEAAASENFLAVKNVRGEYYLNGHWTIEAARALPAASTILHYERGAEGDLAPE 266

Qy        306 TFKIPGPLMADFIFKTRYTAAKDSVVQFFFYQPI       SHQWRQTDFFPCTVTCGGGYQ 365
                   |:      : :      :   |  : :: |:              |    :  |:  || :|
Db        267 RLHARGPTSEPLVIEL ISQEFNPGVHYEYHLPLRRPSPGFSWSHGSWSDCSAECGGGHQ 325

Qy        366 LNSAEC VDIRLFPVVPDHYCHYYPENVKPKPKLKECSMDPCPAPDGTKEIMPYDHFQPL 419
                   | :|         ||| |    |     :|   :  : |:: |||  :      :    ||
Db        326 SRLVFCTID    HEAYPDHMCQRQP   RPADR RSCNLHPCPET      KRTSYLHRPGA 374

Qy        420 PR      -----WEHNPWTACSVSCGGGIQRRSFVCVEESMHGEILQVEEWKCMYAP K 467
               |          |:  |:  || ||||| | ||  |:        |       | | :|
Db        375 WRLAGAQRVCGNSWKAGPWAPCSASCGGGSQSRSVYCISSDGAGIQEAVEEAECAGLPGK 434

Qy        468 PKVMQTCNLFDCPKWIAMEWSQCTVTCGRGLRYRVVLCINHRGE   HVGGCNPQLKLHIK 525
              |  :| |||  |  |     : :|:|:|| |:| | | |   ||    |    |: : : :
Db        435 PPATQACNLQRCAAWSPEPWGECSVSCGVGVRKRSVTCRGERGSLLHTAACSLEDRPPLT 494

Qy        526 EECVIPIPCYKPKEKSPVEAKLPWLKQAQELEETRIATEEPTFIPEPWSACSTTCGPGVQ 585
              | ||            |: :  |            : |            |  | : |  | :
Db        495 EPCV         HEDCPLLSDQAW          HVGT  -     WGLCSKSCSSGTR 527

Qy        586 VREVKCRVLLTFTQTETELPEEECEGPKLPTERPCLLEACDESPASRELDIPLPEDSETT 645
               |:  | :           :   |    |    :|  :| |:  |      ::|   :
Db        528 RRQVICAI            GPPSHC  GSLQHSKPVDVEPCNTQPCHLPQEVPSMQDVH - 573

Qy        646 YDWEYAGFTPCTATCLGGHQEATAVCLHIQTQQTVNDSLCDMVHRPPAMSQACNTEPCP 704
                        || :   :  |  |:  |     |      |          :  |   ::
Db        574           TPASNPWMPGGPQESPASAAPTPATPAVG     - LPAPRLQTQSSRV-  L 616

Qy        705 PRWHVG         SWGPCSATCGVGIQTRDVYCLH PGETPAPPEECRDEK        747
              |||  |         |   |            : |   ||      |  :  :
Db        617 PRWPHGTSRASVARLPWGPLSA        EQVHNTHQPQAQQNEPSECRGSQPGGCYDN 668

Qy        748                        PHALQACNQFDCPPGWH-TEEWQQCSRTCGGGTQ 780
                                     | |  :|    |     |: :      ||: :     ||
Db        669 VATAAGPLGEGCVGQPSHAYPVRCLLPSAHGSC   APWAAPWYTVACVCQGNPPWY   724

Qy        781 NRRVTCPQLLTDGSTLNLSDE       LCQG        PKASSHKSCARTDCPPHLAVGDW 828
                     .    :   :   :  .     |||        |:   :           .     :
Db        725        CH      GNANNPASEQECMSSCQGSLHGPRRPQPGASGRSTHTDGGGSSPACEQ 774

Qy        829 SKCSVSCGVGIQRRKQVCQRLAAKGRRIPLSEMMCRDLPGFPLVRSCQMPECSKIKSEMK 888
                    |   :||                                :  |    |  :   | |   :
Db        775 EPSQHRTGAAVQR                          KPWPSGGLWRQDQQPGPGEAPHTQA 812

Qy        889          TKLGEQGPQI               LSVQRVYIQTREEKRINLTIGSRAYLLPNTS 929
                        :|| : : :               :    | :  :     |     :    :
Db        813 PGEWPWGQELGSRAPGLGPHAPGSPAPPFHSSSYRTSLAGVEPSLVQAALGQ        L 864

Qy        930 VIIKCPVRRFQKSLIQWEKDGRCLQNSKRLGITKSGSLKIHGLAAPDIGVYRCIA    GS 986
              |  : |        :|    : | | : : :| |   :     ||| || : : | | | | : : |
Db        865 VRLSPGSDDIAPESQAAWQKDGQPT SSDRHRLQPDGSLITHPLQAEDAGTYSCGSTRPGR 923

Qy        987 AQETVVLKLIGTDNRLIARPALREPMREYPGMDHGLANSIGVTWHKMRQMWNNKNDLYLD 1046
                : :  |::||      |                    : |      | |
Db        924 DSQKIQLRITG      LCPHP                 THHSHLVSPG                  949

Qy       1047 DDHISNQPFLRALLGHCSNSAGSINSWELKNKQFEAAVKQGAYSMDTAQFDELIRNMSQL 1106
                                                                            |
Db        950                                                            L 950

Qy       1107 METGEVSDDLASQLIYQLVAELAKAQPTHMQWRGIQEETPPAAQLRGETGSVS      QS 1161
              |   |:                 :|  |::|: :     |  |     |  ||   |: :|:         |
Db        951 MIGGD------       --MAVLSEAELS      RPPQPRDP- AQDPGQAGAAGPLGATPS 991

Qy       1162 SHAKNSGKLIFKKPKGPVLMRQSQPPSTSFNKTINSRIGNTVYITKRTEVINTLCDLITPS 1221
              ||  :  :  :|             :  |:||         :  ::: |    :  :| | |          |
Db        992 SHPQPANRLR---     ---LDQNQP      --RVVDASPGQRTRMTCRAEGF          PP 1029

Qy       1222 EATYTWTKDGTLLQP--SVKTILDGTGKTQTQNPTRKEQGIYECSVANHLGSDVESSSVL 1279
               |   | :||   ||  | :  |   | : |         :: | | |   |    |    :
Db       1030 PA IEWQRDG---QPVSSPRHQLQPDGSLVISRVAVEDGGFYTCVAPNGQDRDQRWVQLR 1085

Qy       1280 YAEAPVILSVERNITKPEHNHLSVVVGGIVEAALGANVTIRCPVKGVPQPNITWLKRGGS 1339
                    |   :    :| ||                        |    :  | | |   || | : |
Db       1086 VLGELTISGLPPTVTVPE                        GDTARLLCVVAGESVNIEWSRNGLP 1128

Qy       1340 LSGN    VSLLFNGSLLLQNVSLENEGTYVCIATNALGKAVATSVLHLLERRWPESRIVF 1396
              :   :   |     :|:||:  |:    :||:| | |
Db       1129 VQADGHRVHQSPDGTLLIYNLRARDEGSYTCSA                              1161

Qy       1397 LQGHKKYILQATNTRTNSNDPTGEPPPQLPTWLPGNWSHCSATCGHLGARTQRPQCVMAN 1456
               ||   :  :  ::|    :    |   || :|            :||            |         : :|: :||
Db       1162 YQG-SQAVSRSTLVKVVSPAPTAQPR        PPGR        TR          VDQPE  LAN 1200

Qy       1457 GQEVSEALCD-HLQKPLAGFEPCNIRDCPARWFTSVWSQCSVSC 1499
                        ||   ||   | | |               :::|      |  ||
Db       1201          -CDLTLQAQLCGNE    -         YYSSF     CCASC 1223


RESULT 11
Q9GQR0
ID   Q9GQR0      PRELIMINARY;      PRT;  2174 AA.
AC   Q9GQR0;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Extracellular matrix protein papilin precursor.
GN   PPN OR CG1540 OR CG18436.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DP CN BW;
RX   MEDLINE=20530499; PubMed=10767677;
RA   Kramerova I.A., Kawaguchi N., Nelson R.E., Fessler L.I., Chen Y.,
RA   Kramerov A.A., Kusche-Gullberg M., Kramer J.M., Ackley B.D.,
RA   Sieron A.L., Prockop D.J., Fessler J.H.;
RT   "Papilin in development; a pericellular protein with a homology to the
RT   ADAMTS metalloproteinases.";
RL   Development 127:5475-5485(2000).
DR   EMBL; AF205357; AAG37995.1; -.
DR   HSSP; P12111; 2KNT.
DR   FlyBase; FBgn0003137; Ppn.
DR   InterPro; IPR000561; EGF_like.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR003598; Ig_c2.
DR   InterPro; IPR003600; Ig_like.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR002223; Kunitz_BPTI.
DR   InterPro; IPR000884; TSP1.
DR   InterPro; IPR002221; WAP.
DR   Pfam; PF00047; ig; 2.
DR   Pfam; PF00014; Kunitz_BPTI; 3.
DR   Pfam; PF00090; tsp_1; 5.
DR   Pfam; PF00095; wap; 1.
DR   PRINTS; PR00003; 4DISULPHCORE.
DR   PRINTS; PR00759; BASICPTASE.
DR   ProDom; PD000222; Kunitz_BPTI; 3.
DR   SMART; SM00409; IG; 3.
DR   SMART; SM00408; IGc2; 2.
DR   SMART; SM00410; IG_like; 1.
DR   SMART; SM00131; KU; 3.
DR   SMART; SM00209; TSP1; 7.
DR   SMART; SM00217; WAP; 1.
DR   PROSITE; PS00317; 4_DISULFIDE_CORE; 1.
DR   PROSITE; PS00280; BPTI_KUNITZ_1; 3.
DR   PROSITE; PS50279; BPTI_KUNITZ_2; 3.
DR   PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR   PROSITE; PS50092; TSP1; 3.
KW   Matrix protein; Serine protease inhibitor; Signal.
FT   SIGNAL        1     26       POTENTIAL.
SQ   SEQUENCE   2174 AA;  231936 MW;  038F707952623120 CRC64;

  Query Match             9.6%;  Score 898.5;  DB 5;  Length 2174;
  Best Local Similarity   23.8%;  Pred. No. 5.9e-59;
  Matches  313;  Conservative  164;  Mismatches  467;  Indels  373;  Gaps   48;

Qy         10 VLIGMVFMHSPLPQTIAEKSPGA-YFLPEFALSPQGSFLEDITGEQFLTYRYDDQTSRNI 68
              |:   :      |  |        ::   ||   :|||  :::|  |
Db         17 VLASIHDSQSRFPGLRQKRQYGANMYLPESSVIPGGE----------------------- 53

Qy         69 FSDFDKDGNWLAWCDWSDCSPTCGGGASYSLRRCLTGRN-----CEGQNIRYKICSNHLC 123
               :| |:     |   |  |||||||||| ||   | ||     :         |  |  : || :|:   ||
Db         54 GNDPDE    WTPWSSPSDCSRTCGGGVSYQTRECLRRDDRGEAVCSGGSRRYFSCNTQDC 110

Qy        124 PPDAEDFRAQQCSAYNDVQYQGHYYEWLPRYNDPAAPCALKCHAQGQNLVVELAPKVLDG 183
              |  :   ||||||||  ::    |:  |  :|||:|   |  |   || |  |   :|:         ||:||
Db        111 PEEESDFRAQQCSRFDRQQFDGVFYEWVPYTNAP-NPCELNCMPKGERFYYRQREKVVDG 169

Qy        184 TRCNIDSLDMCISGICQAVGCDRQLGSNAKFDNCGVCAGDGSTCRLVRGQ-SKSHVSPEK 242
              ||||     ||:|::|  |   ||||    |||:||||  |   | ||||||: :|     :   :::|
Db        170 TRCNDELLLVTVNCDGMFVDCDMMLGEDANLLECEKIIGDGSLCKLLENILIIKNLAFCY 229

Qy        243 EEENVIAVPLGSESVEITVECPAHLFIESKTLQGS--KGEHSFNSFGVEVVENTTVEFQ 299
                 :   :: :|  |: ::||     |:  ::   :      |        |:    :  |          |:    :|
Db        230 ND-LLLLPECAINIRIEETVPSSNYLACRNHSCHYYLNGDWKIDFFRPMFFANSWWNYQ 287

Qy        300 RGSFRQTFKIPGPLMADFIFKIRYTAAKDSVVQFFFYQPISHQWRQIDFFPCIVICGGGY 359
              |                     |       |||                            :|||  |
Db        288 R--------------------KPMGFAAPDQ------------------------LTCSGPI 305

Qy        360 QLNSAECVDIRLKKVVPDHYCHYYPENVKPKPKLASCSMDPCPSSDGFKEIMPYDHFQPL 419
                   :     :  ::  |  :    |   |  :    |||::                                 |
Db        306 SESLFIVMLVQEKNISID-YFYSIPESLSHSQ------------------------QDT 339

Qy        420 PRWEHNFWTACSVSCGGGIQRRSFVCVDESMHGEILQVEEWKCMYAPKPKVMQTCNLFDC 479
                 |  |:  :  |||  :||||  |   |        |        :        :  :|      |      ||      |  |         |
Db        340 HTWIHHQFNACSASCGGGNQRRVLCI   NNEILLAVNESLCHQKQKIVLEQACGIEPC 396

Qy        480 -PEWIAMEWSQCTVTCGRGLRYRVVLC  INHEDLHVGGCNPQLKLHIKEECVIPIPCY 535
               |  |:   |||:|:   ||    |   :   |   :  |   |:   ||             ::|:  |      |
Db        397 APHWVEGEWSKCSKGCGSDGFQNRSITCERISSSCEHT----------VEEDAV----CL 442

Qy        536 PRPEPPVLAHLWKLALSEETTELALEEPTLTLWSASLLPCGCVQVKKVKCSVL 595
               |         |                 :  :   |     |           |  :        |:  |      ||   |    |    |:|   |        :
Db        443 KEVGNEP                  AIKQEDNKGVENCPKYHLCPWIPCDKLQCDCKQIRKVICHI 490

Qy        596 TPTQIETELPEEECEGPKLPTERPCLLEACDESPASKELDIPLPEDSETTYDW---EYAG 652
                    :     ||||:|      |      ||:  |||      |:                                     ||      :::|
Db        491 EENGHKEVLPEEDCVEEKFETEKSCLLTPCH                 -GVLWILSQWSS 532

Qy        653 FIPG-----IATCLGGHQEAIAVCLHIQTQQIVNDSLCDMVHRPPAMSQACNIEPCPPRW 707
                  |          |   |    :   |::|                     :     :    |:          |  :|:   |   :     |     :|
Db        533 CNACGGNTPTRTAICGNKEG           KVYPEPPYCE  PEVPTLSEPCKSPKCEAQW 590
```

```
Qy        708 HVCSWGPCSATCGVGIQTRDVYCLHPGE---  --TPAPPE-ECRDE-KPHALQACNQFD- 758
                    |    |||  ||  |:::|  |  |      ||            |||    :  :|    |  ||  :  |  |      :
Db        591 FSSEWSECSAPCGKGVKSRIVIC---CEEDCKIVTPADDSECNKETKPESEQDCRQEKK 637

Qy        759 -CPPGWHIEEWQQCSRTCGGGTQNRRVTCRQLLTDGSFLNLSDELCQGPKASSHKSCART 817
               ||  |      | :||:  ||||  :  |  |  |                              |    |  |
Db        638 VCPGEWFTGPWGKCSKPCGGGERVREVLC-  -------  ---LSNGTK---------- 672

Qy        818 PCPPHLAVCPWSKCSVSCGVGIQRRKQVCQRLAAKGRRIPLSEM----MCRDLPGFPLVR 873
                               ||:|                            :  :   ||||             |  :         ||
Db        673                  SVNCD                      EEKVEPLCEKCNSEACIEDEILPLTS 703

Qy        874 SCQMPECSKIKSEMKTKLGPQGPQILSVQRVYIQTREEKRINLTIGSRAYLLPNTSVIIK 933
              :          |     :   :        | |  ::::|     :        :  |   |:|      ::        |         :::
Db        704 L           LKTLLLLEERELELLLLLTESSTEESSSSSKSLKVEEEECLAKIEIIPEAEDLMQ 756

Qy        934   --CPVRRFQKSLIQWFKDGRGLQNSKRICITKSGSLKIHGLAAPDIGVYRCIAGSAQE 989
                   |     |:  :      :|    |       :::      ||:    ||                                      |    :|
Db        757 SDSPTPYDFPPSTGTTEPSGYDSESTTDSGISTEGS-      ----  -   -GDDEE 798

Qy        990 TVVLKLIGTDNRLIARPALREPMKEYPGMDHSEANSLGVTWHKMKQMWNNKNDLYLDDDH 1049
              |                                      |||::                           ||        |
Db        799 T----------------------  ------  -SEAST---       ---     DLSSSTDS 812

Qy       1050 ICNQPFLRALDDPSNEAGLTNLWLLPNKLLFAAVFQEAYTMLTAQLDLLLKNMSQLMET 1109
                |                 |:  :  |:  |   :     ::       ::|          :  |   |  :                   ||
Db        813 GSTSS------DSISSDSSSSISSDATSEAPASSVSDSSDSIDASP----     ----ET 854

Qy       1110 GFVSDDLASQLIYQLVAELAKAQPTHMQWRGIQEET--          PPAAQLRGETGSVSQS 1161
                 ||||                 |  :  ::  |  :      |   :  |                |  :          :  :      :
Db        855 TGVSDDSTD-   VSSSTEASASESTDVS- GASDSTGSTNASDSTPESSTEASSSTDDST 909

Qy       1162 SHAKNSCKLIPKPKCPVLMKQSQIPSIPPNKIPNSKIINIVYIIKKIEVINILGDLITPS 1221
                :  ||               |       ::    |  |  :  :  :|   |:|   ::   ||                  |
Db        910 DSSDNSSN-  -       -VSESSTEASSSSVSDSSDSSDGSIDGVSSITE           NSS 952

Qy       1222 EATYTWTKDGTLLQPSVKIIIPGTGKIQIQNPT RKEQGIYPCSVANHLGSDVPSSS 1277
                ::|    |  |  |     |         |  |  :      ::  |      |        :  |          |    ||||
Db        953 DSTSDATSDSTASSDSTDSTSDQTTPTTPESSTDSTESSTIDASSTTDASSTSESSS 1009


RESULT 12
Q9VAV4
ID   Q9VAV4      PRELIMINARY,      PRT;   3060 AA.
AC   Q9VAV4; Q9VAV3;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   CG1540 protein.
GN   PPN OR CG1540 OR CG18436.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN BERKELEY;
RX   MEDLINE 20196006; PubMed 10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
CC   -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
CC       SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
DR   EMBL; AE003765; AAF56794.2; -.
DR   EMBL; AE003765; AAF56795.2; -.
DR   HSSP; P12111; 2KNT.
DR   FlyBase; FBgn0003137; Ppn.
DR   InterPro; IPR000561; EGF_like.
DR   InterPro; IPR003598; Ig_c2.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR002223; Kunitz_BPTI.
DR   InterPro; IPR000884; TSP1.
DR   InterPro; IPR002221; WAP.
DR   Pfam; PF00047; ig; 3.
DR   Pfam; PF00014; Kunitz_BPTI; 12.
DR   Pfam; PF00090; tsp_1; 5.
DR   Pfam; PF00095; wap; 1.
DR   PRINTS; PR00759; BASICPTASE.
DR   ProDom; PD000222; Kunitz_BPTI; 12.
DR   SMART; SM00408; IGc2; 3.
DR   SMART; SM00131; KU; 12.
DR   SMART; SM00209; TSP1; 7.
DR   SMART; SM00217; WAP; 1.
DR   PROSITE; PS00317; 4_DISULFIDE_CORE; 1.
DR   PROSITE; PS00280; BPTI_KUNITZ_1; 11.
DR   PROSITE; PS50279; BPTI_KUNITZ_2; 12.
DR   PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR   PROSITE; PS50092; TSP1; 3.
KW   Alternative splicing; Immunoglobulin domain;
KW   Serine protease inhibitor.
FT   VARSPLIC   2803   2803       L -> SVVPV (IN SHORT ISOFORM).
FT   VARSPLIC   2811   2854       INFKIMEISII -> VACFLLHENAV (IN SHORT
FT                                ISOFORM)
FT   VARSPLIC   2855   3060       MISSING (IN SHORT ISOFORM).
SQ   SEQUENCE   3060 AA;  331579 MW;  ACA3ID3EE5807C6 CRC64;

  Query Match             6.5%;  Score 891;  DB 5;  Length 3060;
  Best Local Similarity   23.9%;  Pred. No. 3.8e-58;
  Matches  307;  Conservative  160;  Mismatches  453;  Indels  372;  Gaps   41;

Qy         34 EEPEEALSPQQSFLEDEEGEQFLLYRYDDQISRNIRSDEDKDGNWDAWGDWSDCSRTCGG 93
              :||| :::| |                            :| |:   |   |||||| |:
Db          2 YLPESSVIPGGE                             GNDPGE   WTPWSSPSDCSRTCGG 45

Qy         94 JASYGLRPCLLIPN         TEQNIEYKIGNHGTEEAEEFAQQGAYNEVQYQGHYY 148
              : ||   | ||   :         : :  | ::  ||| :   |||:||| :: ::   :|
Db         46 GVSYQIREGLRREDRCEAVGSGSSRRYESCNIQDCPEEESDFRAQQCSRFDRQQFDGVFY 95

Qy        149 EWLPRYNDPAAPCALKCHAQGQNLVVELAPKVLDGIRCNIDSLDMCISGICQAVGCDRQL 208
              ||:|   | |  '| | | |    ::  :          ||:||| !    !|: ::|  |  ||!!   |
Db         96 EWVPYINAL NGCLLNCMPRGERFYYRQALKVVI   NGRRCEA VNG  EVC DMML 154

Qy        209 GSNAKLDNCGVCAGDGSTCRLVRGQ GKGHVSPLKKELNVIAVPGSGCVRTIVKGPAHL 267
              ||:|||| |  | ||||||: :|    :  ::|    :  : :: :|  : ::|:    |:
Db        155 GSDAKLPKCPRCSGPCSTCKLIRNIITKRLAPCYNL    LLLLPECAINIKIEKIVPSSN 212

Qy        268 FIISKIEQSR     KGERTNGSGIVIVVENIIVLLQRGGERQIIKIPGPLMAPFIEKIRYI 324
              :: :   |      |:  :  :|      |:   :||                    |
Db        213 YLACRNHSGHYYLNGDWRIDFPRPMFFANSWWNYQR                  KPMGF 253

Qy        325 AAKDSVVQFFFYQPISHQWFQIDFFPCIVIGGGGYQLNSAEVIDIRLKFVVPDHYCHYYP 384
              :| |                        :|: |   :     : :: :  : ! | :  |
Db        254 AAPDQ-     --         -     LICSGPISESLFIVMLVQEKNISLD YEYSTP 289

Qy        385 ENVEPRPGERPQLMQFTLQREGFRECMPYIHPQFTTEWFRNEWIACVVCGTSDQRFSFV 444
              |::                                 |      |  |: : ||! ||||| | |
Db        290 ESLSHSQ-     --------     QDTHTWTHHQFNACSASGGGSQSRKVT 324

Qy        445 CVEESMRGLIQQVEEWKCHYAPKPEVMQIPNLEPCTPRWIAMEWSQCTVPCGRGLRYRV 502
              |    :     : :|     |      ||     | |      | | !:  |||:|:  ||  | : |
Db        325 G     NNKTTLALVNEGGGGQKGKEVLEQACTLEPCAPRWVPSEWSKCSKGCGSDGFQNRS 381

Qy        503 VLC   INHRGEHVGGCNPQLKLHIKEECVIPIPCYKPKEKSPVLAKLPWLKQAQELEEIR 560
              : |    |:   |||                 :::: |      | |         |        | :  | |
Db        382 IICERISSSGEHI                 VEEPAV      CLKEVCNKI             AIKQECNF 417

Qy        561 IATEEPTFIPEPWSACSITCGPGVQVREVKCRVLLTFTQTEIELPEEECEGPKLPTERPC 620
                    | :     ||: |     || |  |  | |:| |       :           :  ||||:|      |   ||: |
Db        418 EVRNCPRYHQGPWITDRKLCGRQTRRV     ITIENGHRPVLPEEIPVEERPEIERSC 475

Qy        621 LLEACDESPASRELDIPLPEQSETIYDW      EYAGFTEC          IACTGGHQEATAVCL 672
              ||    |:                            ||     :::|     ·     |      | :·|
Db        476 LLTPCE                      GVEWTTGQWSGCNACGQNIEERIAICGNKEG        512

Qy        673 HIQTQQTVNDSLCDMVHRPPAMSQACNTEPCPPRWHVGSWGPCSATCGVGIQTRDVYCIH 732
                    :    :   |:        :|:  |  :    |     :      |     !|| !|  |::::  | |
Db        513           KVYPEEPCE   PEVPTLSRPCKSPKCEAQWFSSEWSKCSAPCGKGVKSRIVTC      563

Qy        733 PGE          IPAPPE ECRDEPKPHALQACNQFD    CPPGWHIEEWQQCSRTCGGGTQNR 782
               ||           |||  : :|   | ||  :  |  |      :   ||   |        | :||:  ||||  : |
Db        564  GEFDGKIVTPADDDSKCNKETKPESEQPCEGEEKVCPGEWETGPWGKCSKPCGGGERVR 622

Qy        783 RVICRQLLTDGSTLNLSDEL QGPKASSHKSCARHDCPPHLAVGDWSKCSVSCGVGIQRR 842
               | |                     |  . :                                             |||:|
Db        623 EVLC                    LSNGIK                                             GVNCD 637

Qy        843 KQVCQRLAAKGRRIPLSEM         MCRDLPGFPLVRSCQMPECSKIKSEMKTKLGEQGPQI 898
                    :  :  ||||          |  :      :|  :                 |     :  :       | |   :  :
Db        638           EEKVEPLSEKCNSEACTEDLTLPLISI           QKPTEDDSEACDEPGTEL 682

Qy        899 LSVQRVYIQTEEEKRINLTIGSRAYLLPNTGVITK       CPVRRFQKGSTQWEKDGRCLQ 954
              :|     :      :  |  ||:|      ::        |        :::          |      |:  :    :|   |     :
Db        683 IS DGLSDDEKSEDVTDLEGTAKIEITPEAEDLMQSDSPTPYDEPESTGTTEEGSGYTSF 741

Qy        955 NGKKLTTTKPGLKTHSLAAEDTPVYECIAKSALTVVIKETTHNETAPTALEETMRE 1014
              ::     ||:      |                       ,    :||:
Db        742 STTDSGTSTEGS                       SEGEEI                       759

Qy       1015 YPGMDRSEANSIGVIWIIKMRQMWNNKNDLYLDDDRTSNQPELRALLDHPGNDAGSINSWE 1074
                     |||::                  :|    |   |                      |:  :  |:  |  :
Db        760           SEAST                  DLSPSTDSGSTSS         DSTSSDSSSTSSTD 791

Qy       1075 LKNKQFEAAVKQGAYSMPTAQPDELTENMQQMPTTVTTGLAGQTTYQNVAPLAKAQPT 1134
                 ::    ::|    : |  | :                      ||    |:|                   |  :  : |
Db        792 AISEAPASSVSDSSDSTDAS                  TETTGVSDDSTD     VSSSTEASASESST 836

Qy       1135 HMQWRGTQEET               PPAAQTRGETGSVSQSSHAKNSGKTTPKPKGTVTMRQSQPP 1186
                 :    |    :  |               |  ·    .    .     :  :  ||                |      ::
Db        837 DVS    GASDSTGSTNASESTPESSTEASSSTEGSTESSTNSSN           VSESSTEAS 880

Qy       1187 STSPNKTTNGRTGNTVYTTKETEVINITGTTTTPGEATYTWTKPGTTTQPPVKTTTLDGTG 1240
               |  |  :  :  :|    |:|    ::    ||                        |::|      |  |  |          |       | |
Db     887 SSSVSDSNDSSDASTDGVSSTTE [illegible] NSSDSTSDATSDSTASSDSTDSTSDQTT 937

Qy    1247 KIQIQNPT-PKEQGIYECSVANHLGSDVESSS 1277
            :   :: |    |     : |        |   ||||
Db     938 ETTPESSTDSTESSTLDASSTTDASSTSESSS 969


RESULT 13
O76840
ID   O76840      PRELIMINARY;     PRT;  2167 AA.
AC   O76840; Q22911;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   C37C3.6 protein.
GN   C37C3.6.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   Geisel C., Bradshaw H.;
RT   "The sequence of C. elegans cosmid C37C3.";
RL   Submitted (JUL-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- ALTERNATIVE PRODUCTS: TWO FORMS (A AND B) MAY BE PRODUCED BY
CC       ALTERNATIVE SPLICING OF THE SAME GENE. THE SEQUENCE SHOWN IS THAT
CC       OF FORM B.
DR   EMBL; U64857; AAC25868.1; -.
DR   EMBL; U64857; AAC25867.1; -.
DR   HSSP; P00981; 1DTK.
DR   InterPro; IPR003598; Ig_c2.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR002223; Kunitz_BPTI.
DR   InterPro; IPR000884; TSP1.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00014; Kunitz_BPTI; 11.
DR   Pfam; PF00090; tsp_1; 6.
DR   PRINTS; PR00759; BASICPTASE.
DR   ProDom; PD000222; Kunitz_BPTI; 11.
DR   SMART; SM00408; IGc2; 1.
DR   SMART; SM00131; KU; 11.
DR   SMART; SM00209; TSP1; 7
DR   PROSITE; PS00280; BPTI_KUNITZ_1; 10.
DR   PROSITE; PS50279; BPTI_KUNITZ_2; 11.
DR   PROSITE; PS50092; TSP1; 4.
KW   Alternative splicing; Hypothetical protein; Immunoglobulin domain;
KW   Serine protease inhibitor.
FT   VARSPLIC    147    155       MISSING (IN ISOFORM A).
FT   VARSPLIC   1556   1558       KDD -> SKF (IN ISOFORM A).
FT   VARSPLIC   1559   2167       MISSING (IN ISOFORM A).
SQ   SEQUENCE   2167 AA;  237599 MW;  9627178DD52E3635 CRC64;

  Query Match             9.2%;  Score 863;  DB 5;  Length 2167;
  Best Local Similarity   28.6%;  Pred. No. 3e-56;
  Matches  220;  Conservative  114;  Mismatches  312;  Indels  122;  Gaps   30;

Qy      35 LPEFALSPQ[illegible]PTTTPLPT [illegible] TYRYDLQTSENTESDEE 73
           :| :| |     |  |   : :|                       |    | :     :| :
Db      15 VPTWAFSLSSFFGSDVAQKPYIHPNSPPFRDPASSPMKPQAYQVYVDCDVSVTVDKSGQK 74

Qy      74 KI[illegible]NWPAW[illegible]ISKT[illegible]ASY[illegible]LIGENCE[illegible]NIRYKTCSNHDCPPDAEDFRAQ 133
           : || |   ::|||:||||     |:|     :| | :||  :|: : |       ||||:
Db      75 ETGNWGPWVPENECSRSCGGGVQIRKRQC  SCPCT[illegible]ASVRYISCNLNAC ESGTDPRAE 131

Qy     134 QCSAYNDVQY[illegible]YYEW[illegible]AAP[illegible]ALKTHA[illegible]NLVV[illegible]CTECNIL[illegible]M 193
           ||| :||    |::::| | |        | | |  :  |    : | ||:||||:|:: | |:
Db     132 QCSKPNDRAL[illegible]NYHEWTP YKC[illegible]ENKCEIVC[illegible]NFYY[illegible]WALFVVI[illegible]TKCD[illegible]SNII 189

Qy     194 CISCTCQAVGCDRQI[illegible]SNAKPDNCCVCAGDGSTCRLVRCQ-SKSHVSPEKPEENVIAVPL 252
           |: : | ||| :||: | | | || | ||||||: : |:  : :||   :  :| :|
Db     190 CV[illegible] 247
```

```
Qy     253 GSRSVRITVKGPARLFTHSKTLQGFKGHHSPNCPCVFVVENI  TVEPQRGSERQTF 307
           |: ::::|      |       :   |:                |  |:   ||     |:       :| :|
Db     248 GATNIKIQ   EARKSTNNLALKNGPHHPYLNGNGLLQVHEEVEVGGTIPVYDDAEPETL 304

Qy     308 KIPGPLMADPTPKTRY-IAAKDSVVQPFPYQP-- ISHQWRQIDPPISIVICGGGYQLN 362
                |||   :       :     ::|: ::: |  |      : : ::   :: ||:|:|| | |
Db     305 SAQGPLSEELIVALLFRKGSRDIAIKYEESTPLEEEVDYMYKEDNWIPPSVSGGKGVQIR 364

Qy     363 SAECVDIRLKKVVPDHYCHYYPENVKPKPKL-KECSMDPCPSSDGFKEIMPYDHFQPLPR 421
           :  |:| :  |   | |  |      ||    ||:  | |     | :
Db     363 N[illegible]GENK[illegible]EVL[illegible] [illegible]NNA[illegible]VDCEA------ ------- -E 404

Qy     422 WEHNPWTACSVSCGG GIQRRSFVC[illegible]MHGT[illegible]QVEEWKCMYAPKPKVMQTCNLFDCP 480
           |     | :|| :||  | | |    | :    :|   : ||:  |       :| | |||| | ||
Db     405 WPTGPWESCSSTCGDQQQQYRVVYCH[illegible]VPANGRRVTVEDGNCP VERPPVKQTCNRFACP 463

Qy     481 KWIAMEWSQ[illegible]TVTCGPCLPYEVVLCINHRGEHVGGCNPQLKLHIKEECVIPIPCYKPKEK 540
           :| |  || |:  ||    :|| | |     | | |       : ||     : |
Db     464 RWQAGPWSACSEKCGDAPQYRSVTC---RSEKEC---EECKLLAADAC- ------- - 505

Qy     541 SPVEAKIPWIKQAQEIPETRIATEPP----TEIPEPWSACSTTCGPGVQVREVKCRVLLT 596
             | :          : :: :  |        |       ||:      |: | | |        : ||| |:
Db     506 -PAD-------EQEKFDTERTCNLGPCEGLTPVTGEWNLC-TRCNDTEETREVTCK---- 552

Qy     597 [illegible] 656
                       : :|     | |: ||::     | |                |           |:|   :  : : |
Db     553 -----------DSQGRAYPLEK-CLVDNSTEIPTDTRSCATQP---PCEYEWTVSEWSKC 597

Qy     657 TATCLGGHQEAIAVCLHIQT  QQTVNE[illegible]LCDMVHKPPAMSQACNTIE[illegible]KWHVGSWGP 714
           |  |   ||:       :|    |        : |::    |       :|      :      | | |     |:   ||
Db     598 TTECGHGHKIRRVICATHQNGGLEVVPE[illegible]Q-QAEKPEGKINCINEEK[illegible]WYISSWSE 656

Qy     715 CSATCGVGIQTRDVYCLHPGETPAPPEECRDE  KPHALQACNQFDCP 760
           |:| || | | |     ||:    :  |   || | ||   ||     | ||    |||
Db     657 CTAECGGGSQ[illegible]DCP 702


RESULT 14
Q9U8G8
ID   Q9U8G8      PRELIMINARY;      PRT;  3198 AA.
AC   Q9U8G8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Lacunin precursor.
OS   Manduca sexta (Tobacco hawkmoth) (Tobacco hornworm).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC   Sphingiodea; Sphingidae; Sphinginae; Manduca.
OX   NCBI_TaxID=7130;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99457716; PubMed=10528409;
RA   Nardi J.B., Martos R., Walden K.K., Lampe D.J., Robertson H.M.;
RT   "Expression of lacunin, a large multidomain extracellular matrix
RT   protein, accompanies morphogenesis of epithelial monolayers in Manduca
RT   sexta.";
RL   Insect Biochem. Mol. Biol. 29:883-897(1999).
DR   EMBL; AF078161; AAF04457.1; -.
DR   HSSP; P12111; 2KNT.
DR   InterPro; IPR004094; Antistasin.
DR   InterPro; IPR003598; Ig_c2.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR002223; Kunitz_BPTI.
DR   InterPro; IPR000884; TSP1.
DR   InterPro; IPR002221; WAP.
DR   Pfam; PF02822; Antistasin; 4.
DR   Pfam; PF00047; ig; 2.
DR   Pfam; PF00014; Kunitz_BPTI; 10.
DR   Pfam; PF00095; wap; 1.
DR   PRINTS; PR00759; BASICPTASE.
DR   ProDom; PD000222; Kunitz_BPTI; 10.
DR   SMART; SM00408; IGc2; 2.
DR   SMART; SM00131; KU; 10.
DR   SMART; SM00209; TSP1; 7.
DR   SMART; SM00217; WAP; 1.
DR   PROSITE; PS00317; 4_DISULFIDE_CORE; 1.
DR   PROSITE; PS00280; BPTI_KUNITZ_1; 9.
DR   PROSITE; PS50279; BPTI_KUNITZ_2; 10.
DR   PROSITE; PS50092; TSP1; 1.
KW   Immunoglobulin domain; Serine protease inhibitor; Signal.
FT   SIGNAL        1     21       POTENTIAL.
SQ   SEQUENCE   3198 AA;  349364 MW;  AB4ACD459C0D9134 CRC64;

  Query Match              9.2%;  Score 860.5;  DB 5;  Length 3198;
  Best Local Similarity   22.3%;  Pred. No. 8.6e-56;
  Matches  307;  Conservative  191;  Mismatches  503;  Indels  377;  Gaps   56;

Qy         50 IIGEQFLIYRYDDQISRNIESDED            KDGNWDAWGDWSD   CSRICGGGASY 97
                |      |:       ||: : :               |     |||| |     |||||| :
Db         23 IASRHHYIHNVSHRGRHHKQSK HYLSSSYVIPSGSIGWGDWGDSIPCSKIGGGVAS 82

Qy         98 SLRRCI       IGF NCFGQNIRYKICSNHDCPPDAEDFRAQQCSAYNDVQYQGHYYEWIP 152
                | ||         |:  |  |  : :| :|     ||||  :  ||||:|'| |:|:|:    |:: :
Db         83 QKRICLKFIPDGQPQCIGGDIKVFSCPIQDCPPGIDEFAEQCSKYDDVEYRNIKYKWKI 142

Qy        153 RYNDPAAPCALKCHAQQNIVVELAPKVLDGTPCNIDSLDMCISGI QAVI PRQLGGNA 212
                 |        . | |  |  |    |:                |:|||||||  ||  |:|:|:| ||  ||||  ||||.
Db        143  YIRGPNPCELNCMPIIGPPYYEHQLKVIDSTEGNIGSGPEVCVNGI QPVIPGMMLGGNA 201

Qy        213 KEDNGGVCAGDGSIGRLVRGQSKSHVSPEKREENVIAVPLGSRSVRIIVKGPAHLFIESK 272
                :| | |   | |:|  |      |    :         |      ::: :| |: : : |       |: : : :   :
Db        202 REDKCRRCPGNCENCYIINCVLPIQ  DLIRGYNGSLIPEGAIIIVIMEVKIGNNYLALR 260

Qy        273 IIQGS     KGEHSENGPGVPVVENIIVEFQRGGERQTPKIPGEL        MADPIFKIRYI 324
                  |:     |::  : |    : :            ::|      | |  |  |  .         : : : :
Db        261 AKNGIYYLNGDYHILPPRSMMIAGALWEYER   SQQGFAAPDKLRCIGPIIEPLYLSLLL 318

Qy        325 AAKDGVVQFFFYQPI               SHQWRQIDFFPCIVIGGGGYQLNSAEC     VGIRLKK 373
                 : :  : : : :  |              : : |      |   :  :|||:|:|   :   .     :|: |
Db        319 QSVNVGIFYEYSLPIIRAPFPIQNYIWVHEHEIEGSAICGGGFQIRNVIGRIQMGLEI      376

Qy        374 VVPDHYCHYYFFNVKPKPKIKPCSMLPCPSSGCFKKIMFYDHFQPLPRWEHNPWIACSVS 433
                 |  |   |        |  :||        :  | :  :  ||                    |:|      |  :||
Db        377  VDDIIC     DEGLKPVIN QICNNNVCP                        PQWVIGEWGSCSHP 414

Qy        434 CGG GIQRRSEVCVEESMPDLLLQVEDWKC    MYAPKPKVMQICNL PICPRWIAMEWSQ 489
                || |  |  |     |  :  ::|        |  :  :|   :   ||| :    :  |||:      |||    |      |
Db        415 CGSDGIQSRRIQCEKIIVNGERSIVSDKECFHLDCPKPEKPRKCNVNIICPIWEICPWKP 474

Qy        490 CIVTCGRGLRYRVVLCINHRGEHVGGCNPQLKLHIKKECVIPIP YAPKKKSIVEAKLPW 549
                :     | |  :  | |:|          :    :                ||       |
Db        475 CDIIYGEGKQIRQVVCYQKNGRRIDVLD          DSECIDERP                     511

Qy        550 EKQAQELEEIRIAIEEPI  EIPEPWSACSIPCGPGVQVREVKCRVLLIPIQIEIELPEEE 608
                   :  ::    |  : : |   :  :     || |     |          :: :  |     ||
Db        512      ESEQKCMIHSEERIDWVASEWSGCD NCESIMKIRIAKC        -     -     -   549

Qy        609 CEGPKIPIEWPCIIPACDPCPASRPIDIPCDPGQCIIYIWPYACPIPCTAPCCCHQPAC 670
                                                               ||||
Db        550                                                  IIYD             --     553

Qy        669 AVCHEQIIAQIVNECLCNMVGHE   IAMGQACNI      GECIIRWHVGWGR  SAP IVSIQ 724
                         :: |:| |  |      .  |  |  :  :  |:          ||      :|:      |     ||: | |  |  |
Db        554              RKLVDIGSEC    AHHPIPVLQEPCDAAILPPCDVQWYAIQWSKCSSECGDGIQ 603

Qy        725 IRDVYC LHPGEIPAPPEECR               DEKPHALQACNQFDCPPGWHIEEWQQCSRICG 776
                || |:| ::        |       ||  :                ||  :           :    |    |:      |  : | : | |
Db        604 TRRVFCGIYVNNIVVEVEEQRCAGIPKYGPKENC    IVIEEKCPIEWYIAPWSEQSKPCG 661

Qy        777 CGIQNRRVIC    RQLLIDGSPLNLSD   ELCQ GP            KASSHKS   CARIDCPP 821
                :|  |:||||  |       |  ::|          .  |   : :  |  ||                        ||   :              |.
Db        662 CGEQSRRVICLRGDQEVSEPCGINVPESVQVCNICPCNKPEIIIIRGHCIPIMEEEECC           719

Qy        822 HLAVGDWSKCSVSCGVGIQRRKQVCQRLAAKGRKIPLSEMMCRDLPGPPLVRSCQMPEGS 881
```

```
                |:  :            |:                     ||   ||    |    ::|     .   |      .
Db        720            DYEEYYIEIGLDF                 ALAFSGRISE   IGMISEIISALG               754

Qy        882 KIKSEMKIKDGPQGPQILSVQRVYIQIREEKRINIIGSRA                          YLI 925
                            : : :      |:        ||         ::     |:         || :                     |
Db        755                   DSGIEIGSIGDYIIVSDDGGIS  DDGSGSGFSSEGIDFISDIIDDYSI 802

Qy        926 PNISVIIKCPVERPQKGLIQWLKDGRCLQNGKRIGIIKGGSLKIHGLAAPDGVYRCIAG 985
                 :|       :                :  |:              :  :   |          : : :| : :        .         :
Db        803 DSIDSIVT                ESSL          SPSSVSDLKHIPRSENGIVKPKAAIAI          KDIIP 846

Qy        986 SAQEIVVLKL          -                                    IGIDNKLIARPALREPMKEYPGMDH    S 1021
                 |   |  |   |                                                    |:|::                    ::      |      :
Db        847 SEIEISVIDLDYSIIEDIDYIVIDSSISSDEIIGIDDK             SGIIIDPSSSDSIII 899

Qy       1022 EANGEGVIWHKMRQMWNNKNIIYIDDGHIGNQPIIRALLGHCNGAGINCWIIKNKQFE 1081
                |        |:|                    :  :  |           :     : :  |   |                 |  : :|         |           |:|
Db        900 EIYESGLI                          DSSILKEISEIAIDIPGL                   SIISGDEISASDKSK          940

Qy       1082 AAVKQGAYGMDIAQPHGLIENMQQLMLPGEVSPPLASQLIYQIVAEIAKAQ  PIHMQWKG 1140
                              |  |  : :    |  :  :   |            |       :|:          :              |     |          |        :  :     |
Db        941                 SIIGSEIIEKLIGISYSAIIESSIDESIVEEESSGYIIDISASISGLLSSRD 992

Qy       1141 IQEE    IPPAAQLRGEIGSVSQSSHAKNSGKLIPKPKGPVIMRQSQPPSISPNKIINSKI 1198
                 |  :|      ||             :    |     |    |    |    |     :                          ||     :        :      :     : :    :  : :     :|      :   |
Db        993 IIDESSIPYIKDIEGIIASGSIES              IPIDE  IIMSKVIEESSVAEPEIIKIII 1043

Qy       1199 GNIVYIIKPIEVIN-          -                                     IIGDLIIPSEAIYIWIKDGIIIQPS 1237
                         |        |        :        |                                                             |     : :               :       :      |            :|!     :     :
Db       1044 TEEVSGISESASINSDKIIMIIISEDIGKISVSEIIIEMIVIEEISEISPIEGISDKII 1103

Qy       1238 VKIIIIGPGKIQIQNGIPEEQGIYEGVANHGGIVEGGGVLYARAPVIISVEENIIK 1295
                     :      :      :     |              :          |      !      :     |     :                    :!:    |:                    ::                :     : :                        |
Db       1104 MSIVSEEIESSSVIEEIIIEIIVVENA               IISSIIEVIASGKIIMIIMSEESEK 1155


RESULT 15
O44938
ID   O44938      PRELIMINARY;      PRT;  1572 AA.
AC   O44938;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Thrombospondin.
GN   THR1.
OS   Haemonchus contortus (Barber pole worm).
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;
OC   Trichostrongyloidea; Haemonchidae; Haemonchinae; Haemonchus.
OX   NCBI_TaxID=6289;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MOREDUN;
RA   Skuce P.J., Newlands G.F.J., Stewart M., Pettit D., Smith D.,
RA   Knox D.P.;
RT   "Cloning and characterization of thrombospondin, a novel multidomain
RT   glycoprotein associated with the gut of Haemonchus contortus.";
RL   Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF043121; AAB99830.2; -.
DR   HSSP; P05067; 1CA0.
DR   InterPro; IPR000561; EGF_like.
DR   InterPro; IPR002223; Kunitz_BPTI.
DR   InterPro; IPR000884; TSP1.
DR   Pfam; PF00014; Kunitz_BPTI; 6.
DR   Pfam; PF00090; tsp_1; 6.
DR   PRINTS; PR00759; BASICPTASE.
DR   ProDom; PD000222; Kunitz_BPTI; 6.
DR   SMART; SM00001; EGF_like; 1.
DR   SMART; SM00131; KU; 6.
DR   SMART; SM00209; TSP1; 7.
DR   PROSITE; PS00280; BPTI_KUNITZ_1; 6.
DR   PROSITE; PS50279; BPTI_KUNITZ_2; 6.
DR   PROSITE; PS50092; TSP1; 2.
KW   Serine protease inhibitor.
```

```
SQ   SEQUENCE   1572 AA;  171871 MW;  2260B30DC2F903EC CRC64;

  Query Match             8.8%;  Score 825.5;  DB 5;  Length 1572;
  Best Local Similarity   28.8%;  Pred. No. 1.3e-53;
  Matches  207;  Conservative  107;  Mismatches  286;  Indels  119;  Gaps   30;

Qy         73 DKDCNWDA--WCDW--SDCSRTCCCCASYSLPPCLTCRNCPCQNIRYKTCSNHDCPPDAE 128
              || |  ::  || |     |||||||     |:|     :| | ::|| :|:    | |:
Db         65 DKTGQSESGPWGPWVPEQCSRTCCCGVQTEKRQC--PCDCTCPSVRYVSCNLEPC-PEGT 121

Qy        129 DFRAQQCSAYNDVQYQGHYYEWLPRYNDPAAPCALKCHAQGQNLVVELAPKVLDGIRCNT 188
              ||||:||:|:||     | |::|:| |        | | |  :   |    :     |:|||:|::
Db        122 DFRAEQCAAHNDLFILDQYHEWIP-YEG-ENPCELTCRPENGNEYYEWDLIVVLGIEGDS 179

Qy        189 DSLDMCISGICQAVGCDRQLGSNAKEDNCGVCAGDGSTCRLVRCQ-SKSHVSPEKREENV 247
                 |:|: |:|   :|||  :|:|    | | |||| |||| |: : |     : ::||    :  |
Db        180 KCDHICVDCVCLPLCCPCKLCSALKLDKCGVCPGDGSQCKITECSPDERNLSPCYHD--V 237

Qy        248 IAVPLGSRSVRITVKGPAHLFIESKTLQGSKCEHSFNSPCVFVVENTT------VEFQRG 301
              : :| |: ::||      ::       :  :: |    :   |    :  ||           |:  |
Db        238 MKIPAGATAIRIEEARPSS---NNLAMKNSSDYYFLNGNSMIQVEKDVELNGVHFEYDDC 294

Qy        302 S-ERQTFKIPGPLMADFIFKIRYTAA-KDSVVQFFFYQPI----SHQWRQTDFFPCIVIC 355
                 || | |   |||  : :                :|   ::: |    ||        : ::   ::  |:|||
Db        295 KFFRITAK--GPLFFFVVVSVLIPKGNREVSIKYEFSVPIVDEVDYMYKPGEWSACSVTC 352

Qy        356 GGGYQLNCAPCVLIKIEKVVIHGYCHYSIFNVKIKIKLEKISMKPIPSSDGFKEIMPYDH 415
              | | |  :  |: : :  : | | |     :     ||: ::     || : |   |
Db        353 GKGQQTRTIPYCIETKTQARVDDQLC---DDANSIKPEFEK----PCEIVDCQAE------ 399

Qy        416 FQPLPRWEHNPWIACSVSCG-GGIQRRSFVCVEESMHGEILQVEEWKCMYAPKPKVMQTC 474
                    |    |   || :||  | | |   | :   :|   : |::  |  | :| | | |
Db        400 ------WEKGPWEPCSQTCGPQGFQYRVVYCHQVEANGRRIIVDDGNCT-AERPAVRQVC 452

Qy        475 NLFDCPKWIAMEWSQCTVTCGRGLRYRVVLCINHRGEHVG------GCNPQLKLHIKEEC 528
              | | ||:| |  || |:  ||   :| | | : :    |       |: :  |  :  |
Db        453 NRFACPEWQAGPWSACSEKCGDAFQYRSVICKSEKEGEEGKLLPAEACSSENILESQRSC 512

Qy        529 VIPIPCYKPKEKSPVEAKLPWLKQAQELEETRIATEEPTFIPEPWSACSTTCGPGVQVRE 588
               :  ||   |                             |    |  ||   |    : ||
Db        513 NLG-PCTGLK-----------------------------FFTSEWKLCG-KCNDTEETRE 541

Qy        589 VKCRVLLTFTQTETELPEEECEGPKLPTERPCLLEACDESPASRELDIPLPEDSETT--- 645
              | |:                  : :|   | |: || :                | :||:   :  |
Db        542 VTCK---------------DSQGRAYPLEK-CLTD--------DEKEIPVDIRACATQQP 577

Qy        646 --YDWEYACFTPCTAICLGGHQEAIAVC-LHIQTQQIVNDSLCDMVHRPPAMSQACNTEP 702
                |:|   : :: |:   |   ||:      :| :| :    || |           :|         | |
Db        578 CNYEWTASQWSKCSIECGHGHKTRRVICAIHEECDVIVVDEALCQGEKPEVKMNCTNEEK 637

Qy        703 CPPRWHVGSWGPCSATCGVGIQTRDVYCLHPGETPAPPEEC-RDEKPHALQACNQFDCP 760
              |    |: | | |||  || | | | |    ||:  :  |  || |    | |: :| ||  :||
Db        638 CTGTWYSGPWSPCSVECGGGKQERVAVCLNYDKKPV-PEWCDESEMPNLIQDCNTDECP 695

Search completed: March 22, 2003, 13:30:44
Job time : 86 secs
```

```
                      GenCore version 5.1.4_p5_4578
              Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on:          March 22, 2003, 13:27:54 ; Search time 25 Seconds
                                            (without alignments)
                                            1990.166 Million cell updates/sec

Title:           US-09-784-358-2
Perfect score:   9346
Sequence:        1 MASWISPWWVLIGMVFMHSP..........LNDCSLDRYKQKCCQSCQES 1693

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        262574 seqs, 20422922 residues

Total number of hits satisfying chosen parameters:       262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Issued_Patents_AA:*
                 1:  /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
                 2:  /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
                 3:  /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
                 4:  /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
                 5:  /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
                 6:  /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                   SUMMARIES
                      %
Result              Query
  No.    Score      Match  Length DB  ID                         Description

    1     2048      21.9     525  4  US-09-369-364A-21          Sequence 21, Appl
    2     1028      11.0    1382  4  US-09-369-364A-13          Sequence 13, Appl
    3    812.5       8.7    1224  4  US-09-930-872-4            Sequence 4, Appli
    4    798.5       8.5    1081  4  US-09-369-364A-17          Sequence 17, Appl
    5    595.5       6.4    1205  4  US-09-491-522-11           Sequence 11, Appl
    6    571.5       6.1     907  4  US-09-130-491-2            Sequence 2, Appli
    7    564.5       6.0     591  4  US-09-130-491-16           Sequence 16, Appl
    8    553.5       5.9    1211  4  US-09-491-522-5            Sequence 5, Appli
    9      545       5.8     450  4  US-09-369-364A-19          Sequence 19, Appl
   10      519       5.5     855  4  US-09-369-364A-5           Sequence 5, Appli
   11    506.5       5.4     874  4  US-09-369-364A-15          Sequence 15, Appl
   12    499.5       5.3     481  4  US-09-130-491-8            Sequence 8, Appli
   13    495.5       5.3     608  4  US-09-130-491-13           Sequence 13, Appl
   14    484.5       5.2     905  4  US-09-369-364A-9           Sequence 9, Appli
   15      482       5.2     930  4  US-09-369-364A-2           Sequence 2, Appli
   16      477       5.1     967  4  US-09-122-126B-15          Sequence 15, Appl
   17      474       5.1     528  4  US-09-369-364A-22          Sequence 22, Appl
   18      431       4.6     907  4  US-09-369-364A-7           Sequence 7, Appli
   19      415       4.4     837  4  US-09-122-126B-2           Sequence 2, Appli
   20    365.5       3.9     411  3  US-08-985-526-3            Sequence 3, Appli
   21    341.5       3.7     768  2  US-08-918-914-4            Sequence 4, Appli
   22      273       2.9     459  1  US-08-313-288B-15          Sequence 15, Appl
   23      265       2.8     807  1  US-07-862-021B-10          Sequence 10, Appl
   24      265       2.8     807  1  US-08-313-288B-10          Sequence 10, Appl
   25      265       2.8     807  5  PCT-US93-03164-10          Sequence 10, Appl
   26      264       2.8     568  1  US-07-862-021B-14          Sequence 14, Appl
   27      264       2.8     568  5  PCT-US93-03164-14          Sequence 14, Appl
   28      256       2.7     802  1  US-07-862-021B-12          Sequence 12, Appl
   29      256       2.7     802  1  US-08-313-288B-12          Sequence 12, Appl
   30      256       2.7     802  5  PCT-US93-03164-12          Sequence 12, Appl
   31      240       2.6    1170  1  US-08-313-288B-20          Sequence 20, Appl
   32      240       2.6    1172  1  US-08-313-288B-19          Sequence 19, Appl
   33    213.5       2.3     630  2  US-08-752-907B-14          Sequence 14, Appl
   34    213.5       2.3     630  4  US-09-707-802-14           Sequence 14, Appl
   35    213.5       2.3     630  4  US-09-991-329-14           Sequence 14, Appl
   36      207       2.2     607  2  US-08-752-907B-12          Sequence 12, Appl
   37      207       2.2     607  4  US-09-707-802-12           Sequence 12, Appl
   38      207       2.2     607  4  US-09-991-329-12           Sequence 12, Appl
   39      203       2.2     805  3  US-08-985-526-34           Sequence 34, Appl
   40      202       2.2     806  2  US-08-443-861-5            Sequence 5, Appli
   41      202       2.2     806  4  US-08-153-829B-5           Sequence 5, Appli
   42      202       2.2    1367  1  US-07-813-593-4            Sequence 4, Appli
   43      202       2.2    1367  1  US-07-977-451-6            Sequence 6, Appli
   44      202       2.2    1367  1  US-07-946-507-4            Sequence 4, Appli
   45      202       2.2    1367  1  US-08-252-517-6            Sequence 6, Appli


                                  ALIGNMENTS


RESULT 1
US-09-369-364A-21
; Sequence 21, Application US/09369364A
; Patent No. 6391610
; GENERAL INFORMATION:
;  APPLICANT:  Apte, Suneel
;  APPLICANT:  Hurskainen, Tiina L.
;  APPLICANT:  Hirohata, Satoshi
;  TITLE OF INVENTION:  Nucleic Acids Encoding Zinc Metalloproteases
;  FILE REFERENCE:  [illegible]
;  CURRENT APPLICATION NUMBER: US/09/369,364A
;  CURRENT FILING DATE:  1999-08-06
;  NUMBER OF SEQ ID NOS:  [illegible]
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 21
;   LENGTH: 525
;   TYPE: PRT
;   ORGANISM: Homo sapiens ADAMTS-R1
US-09-369-364A-21

  Query Match             21.9%;  Score 2048;  DB 4;  Length 525;
  Best Local Similarity   68.3%;  Pred. No. 8.3e-148;
  Matches  344;  Conservative   72;  Mismatches   88;  Indels    0;  Gaps    0;

Qy         64 ISRNIKSGDLSKPGNWDAWGPWSECSRTCGGGASYSLRRCLTGRNCEGQNIRYKTCSNHDC 123
              :|:  ||:|||:|| ||||| ||:|||||||||:: |||::|||| :::|||:.| ::||| ||
Db         22 SSRIARSEEDRDGLWDAWGPWSECSRTCGGGAANSLRRCLSSKSCEGRNIRYRTCSNVDC 81

Qy        124 PPDAEDFRAQQCSAYNDVQYQGHYYEWLPRYNDPAAPCALKCHAQGQNLVVELAPKVLDG 183
              ||:| |||||||||:|:: | :|:|:|  ||| :|: | ::|  |: |||||||||||||
Db         82 PPEAGDFRAQQCSAHNDVKHHGQFYEWLPVSNDPDNPCSLKCQAKGTTLVVELAPKVLDG 141

Qy        184 TRCNTDSLDMCISGICQAVGCDRQLGSNAKEDNCGVCAGDGSTCRLVRGQSKSHVSPEKR 243
              ||| |:|||||||||:|| ||:| ||||  :|||||||| ||||||||||||| :| :|  |
Db        142 TRCYTESLDMCISGLCQIVGCDHQLGSTVKEDNCGVCNGDGSTCRLVRGQYKSQLSAIKS 201

Qy        244 EENVIAVPLGSRSVRITVKGPAHLFIESKTLQGSKGEHSFNSPGVFVVENTTVEFQRGSD 303
              :: |:|:|  ||: :|: :||| |:::|:|||:|:||:|::|  :| |:|::|:|:|: :
Db        202 DDTVVAIPYGSRHIRLVLKGPDHLYLETKTLQGTKGENSLSSTGTFLVDNSSVDFQKFPD 261

Qy        304 KQTFKIPGPLMADFIFKTRYTAAKDSVVQFFFYQPISHQWRQTDFFPCTVTCGGGYQLNS 363
              |: :: :||| :|||| | :: :  || ||| ||||| |:|::|||||:: ||||||||| |
Db        262 KEILRMAGPLTADFIVKIRNSGSADSTVQFIFYQPIIHRWRETDFFPCSATCGGGYQLTS 321

Qy        364 AECVDIRLKRVVPDHYCHYYPENVKPKPKLKECSMDPCPSSDGFKEIMPYDHFQPLPRWE 423
              ||| |:|  ||| :|:|||||||:||||||:||::||||:|||:|:|||||  :||||||
Db        322 AECYDLRSNRVVADQYCHYYPENIKPKPKLQECNLDPCPASDGYKQIMPYDLYHPLPRWE 381

Qy        424 HNPWTACSVSCGGGIQRRSFVCVEESMHGEILQVEEWKCMYAPKPKVMQTCNLFDCPKWI 483
                |||||| :|:||||  ::  |||| :  | :  ||| |||| ||  :  | ||:|||||:
```

```
Db     382 AIPWIAQSSSGPPIQEEAVSCVFEEIQCHVISVEEWECMYLPKMPIAQPCNIFDCPKWL 441

Qy     484 AMEWSQCTVTCGRGLRYRVVLCINHRGEHVGGCNPQLKLHIKEECVIPIPCYKPKEKSPV 543
           | ||| |||||:||||||||||:||| | |||::: | |||||||::| ||||||| ||
Db     442 AQFWSPCTVTCGQGLRYRVVLCIDHRGMHTGGCSPKTKPHIKEECIVPTPCYKPKEKLPV 501

Qy     544 EAKLPWLKQAQELEETRIATEEPT 567
           |||||| ||||||||     :|||:
Db     502 EAKLPWFKQAQELEEGAAVSEEPS 525

RESULT 2
US-09-369-364A-13
; Sequence 13, Application US/09369364A
; Patent No. 6391610
; GENERAL INFORMATION:
;  APPLICANT:  Apte, Suneel
;  APPLICANT:  Hurskainen, Tiina L.
;  APPLICANT:  Hirohata, Satoshi
;  TITLE OF INVENTION:  Nucleic Acids Encoding Zinc Metalloproteases
;  FILE REFERENCE:  26473/4007/10-30-00
;  CURRENT APPLICATION NUMBER: US/09/369,364A
;  CURRENT FILING DATE:  1999-08-06
;  NUMBER OF SEQ ID NOS:  31
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 13
;   LENGTH: 1882
;   TYPE: PRT
;   ORGANISM: Homo sapiens ADAMTS-9
;   FEATURE:
;   NAME/KEY: MOD_RES
;   LOCATION: (468)
;   OTHER INFORMATION: Xaa = C
;   NAME/KEY: MOD_RES
;   LOCATION: (521)
;   OTHER INFORMATION: Xaa = Y
US-09-369-364A-13

  Query Match             11.0%;  Score 1028;  DB 4;  Length 1882;
  Best Local Similarity   28.4%;  Pred. No. 3.7e-69;
  Matches  353;  Conservative  168;  Mismatches  456;  Indels  750;  Gaps   65;

Qy      75 DCNWDAWGDWSECSRTCGGGASYSLRRCLT------GRNCEGQNIRYKTCSNHDCPPDAE 128
           ||:| :|   :  |||||||    ::| |          |: | |: :::|:|:     |
Db     535 DGSWGSWSPFGTCSRTCGGGIKTAIRECNRPEPKNGGKYCVGRRMKFKSCNTEPCLKQKR 594

Qy     129 DFRAQQCSAYNDVQYQGHYY              FWIPRYNDPAAP--CALKCHAQGQNLVVELA 177
           ||| :||:        : | ::          |:|:|:         | | |   |    :|
Db     595 DFRDEQCA-----HFDGKHFNINGLLPNVRWVPKYSGILMKDRCKLFCRVAGNTAYYQLR 649

Qy     178 PKVLDGTRCNTDSLDMCISGICQAVGCDRQLGSNAKEDNCGVCAGDGSTCRLVRGQSKS  236
            :|:||| |   |: |:|: |:|:  |||  | | |: | |||| || |:|: | |    :
Db     650 DRVIDGTPCGQDTNDICVQGLCRQAGCDHVLNSKARRDKCGVCGGDNSSCKTVAGTFNTV 709

Qy     237 HVSPEKREENVIAVPLGSRSVRITVKGPAHLFIESK-----TLQGSKGEHSFNSPGVFV 290
           |        :: :| |: :: :       | |         |   ||||    |   | ||
Db     710 HYG----YNTVVRIPAGATNIDVR    QHSFSGETDDDNYLALSSSKGEFLLN  GNFV 759

Qy     291 VE-------NTTVEFQRGSE------RQTFKIPGPLMADFIF-------KTRYT---A 325
           |          |  ||:  |||          | :|   |:   :           ||:
Db     760 VTMAKREIRIGNAVVEYS GSETAVERINSTDRIEQELLLQVLSVGKLYNPDVRYSFNIP 818

Qy     326 AKDSVVQFFF ----YQPISH---------------------------------- 341
            :|   ||::      :|  |
Db     819 IEDKPQQFYWNSHGPWQACSKPCQGERKRKLVCTRESDQLTVSDQRCDRLPQPGHITEPC 878

Qy     342       QWRQTDFFPCTVTCGGGYQLNSAEC-----VDIRLKRVVPDHYCHYYPENVKPK 390
                 :|          |:  || ||:      |      :| : :: | | :|  :      ||
Db     879 GTGCDLRWHVASRSECSAQCGLGYRTLDIYCAKYSRLDGKTEK-VDDGFCSSH---PK 932

Qy     391 PKLKECSMDPCPSSDGFKEIMPYDHFQPLPRWEHNPWTACSVSCGGGIQRRSFVCVEESM 450
           |  :|      | :                  | :: || || || || |||  :||   :
Db     933 PSNREKCSGECNTGG----- -----------WRYSAWTECSKSCDGGTQRRRAICV--NT 974
```

```
Qy     451 HGEILQVEEWKCMYAPKPKVMQTCNLFDCPKWIAMEWSQCTVTCGRGLRYRVVLC----- 505
              ::|  :: || :  |  : | |: | ||:| : :||:| ||||:| :: | |
Db     975 RNDVL--DDSKCTHQEKVTI-QRCSEFPCPQWKSGDWSECLVTCGKGHKHRQVWCQFGED 1031

Qy     506  INHRGEHVGGCNPQLKLHIKEECVIPIPCYKPKEKSPVEAKLPWLKQAQELEETRIATE 564
            :| |        |:|: |      : | |                              |
Db    1032 RLNDR-----MCDPETKPTSMQTCQQP------------------------------E 1054

Qy     565 EPTFIPEPWSACSTTCGPGVQVREVKCRVLLTFTQTETELPEEECEGPKLPTE-RPCLLE 623
             ::    ||   || ||| | |:| ||| :: |:              : : :|      ||: : | |
Db    1055 CASWQAGPWVQCSVTCGQGYQLRAVKC-IIGTYMSV      VDDNDCNAATRPTDTQDCELP 1110

Qy     624 ACDEEPASRELDIPLPEDSETTY        DWEYAGFTPCTATCLGGHQEAIAVCLHIQTQQ 678
           :|   ||:         ||   :||       | :  :|||:|||   | :      |   :
Db    1111 SCHPPPAA-------PETRRSTYSAPRTQWRFGSWTPCSATCGKGTRMRYVSC-RDENGS 1162

Qy     679 TVNDSLCDMVHRPPAMSQACNTEPCPPRWHVGSWGPCSATCGVGIQTRDVYCLHPGETPA 738
              ::| |  : || |   : |:  ||   :|       |  || ||| |   || | |::   :
Db    1163 VADESACATLPRPVA-KEECSVTPC-GQWKALDWSSCSVTCGQGRATRQVMCVNYSDHVI 1220

Qy     739 PPEEC-RDEKPHALQACNQFDCPP-------------------------- GWH 764
              || :|  |   | |:    ||                                   |
Db    1221 DRSECDQDYIPETDQDCSMSPCPQRTPDSGLAQHPFQNEDYRPRSASPSRTHVLGGNQWR 1280

Qy     765 IEEWQQCSRTCGGGTQNRRVTCRQLLTDGSFLNLSDELCQGPKASSHKSCARTDCPPHLA 824
               |  || || ||:| | | |:      :|   |   :  :  |     ::|    | | |
Db    1281 TGPWGACSSTCAGGSQRRVVVCQD  ENGYTAN---DCVERIKPDEQRACESGPC-PQWA 1334

Qy     825 VGNWSKCSVSCGVGIQRRKQVCQRLAAKGRRIPLSEMMCRDLPGLPLVRSCQMPECSKIK 884
             |:| :|:  || ||: |   ||| :  | | |  ::  |   |   |    |    |
Db    1335 YGNWGECTKLCGGGIRTRLVVSQR--SNGERFP--DLSCEILDKPPDREQCNTHAC---P 1387

Qy     885 SEMKTKLGEQGPQILSVQRVYIQTREEKRINLTIGSRAYLLPNTSVIIKCPVRRFQKSLI 944
            :    |                                     | :|  : |     |:::
Db    1388 HDAAWSTG                -------------PWSSCSVSCGRGHKQRNVY 1414

Qy     945 QWEKDGRCLQNSKRLGITKSGSLKIHGLAAPDIGVYRCIAGSAQETVVLKLIGTDNRLIA 1004
              |||  |::                          |                 : :|
Db    1415 CMAKDGSHLESD-------------- ---------Y------------ KHLA 1432

Qy    1005 RPALREPMREYPGMDHSEANSLGVTWHKMRQMWNNKNDLYLDDDHISNQPFLRALLGHCS 1064
           :|                                    |          ||
Db    1433 KP-------------------- -------------------H-------- ---GHRK 1439

Qy    1065 NSAGSTNSWELKNKQFEAAVKQGAYSMDTAQFDELIRNMSQLMETGEVSDDLASQLIYQL 1124
             |     |                   | ||:|
Db    1440 CRGGRCPKW---------- --KAGAWS------------------------- - 1454

Qy    1125 VAELAKAQPTHMQWRGIQEETTPAAQLRGETGSVSQSSHAPNSGKLTFKPKGPVLMRQSQ 1184
                  | :      ||:|:
Db    1455 -------QCSVSMGRGVQQ---------------- ------------ -- ----- 1466

Qy    1185 PPSISFNKTINSRIGNTVYITKRTEVINILCDLITPSEATYTWTKDGTLLQPSVKIILDG 1244
                  : :   :|| |   | : ||       |
Db    1467 -------RHVGCQIG-THKIARETE - --C---------------- -- -    - -- 1484

Qy    1245 TGKIQIQNPTRKEQGIYECSVANHLGSDVESSSVLYAEAPVILSVERNITKPEHNHLSVV 1304
                  ||   :  :   ||                        :|
Db    1485 --------NPYTRPESECEC-- --------------QGP--  - ------- -  - -- 1499

Qy    1305 VGGIVEAALGANVTIRCPVKGVPQPNITWLKRGGSLSGNVSLLFNGSLLLQNVSLENEGT 1364
                            |||:
Db    1500 ---------------RCPL---------- --- -------- ----------    - - 1503

Qy    1365 YVCIATNALGKAVATSVLHLLERRWPESRIVFLQGHKKYILQATNTRTNSNDPTGEPPPQ 1424
           |
Db    1504 YT---------------------------------------- -  - -  - - -- 1505

Qy    1425 EPFWEPGNWSHCSATCGHLGARIQRPQCVMANGQEVSEALCDHLQKPLAGFEPCNIRDCP 1484
              |     |    |: |||   |:|  ::    ||    |    ||   | ||    ::|:     | |:::  |
Db    1506 ---WRAFEWQECTKTCGE-GSRYRKVVCVDDNKNEVHGARCDVSKRPV-DRESCSLQPCE 1560
```

```
Qy       1485 AFWFLSVWSQCSVSCGEGYHSRQVTCKRT     KANGTVQVVSPRA APKDRPLGRKPCFG 1541
Db       1561 YVWITGEWSECSVTCGKGYKQRLVSCSEIYTGKENYEYSYQTTINCP        GTQPPSV 1614

Qy       1542 HPCVQWEPGNWSECSKTCGRAVEMQLRHTACQHNSSGQNCTGRKETTKRNCTSSACDVCW 1601
Db       1615 HPCYLRE    CP                                          VSAIW 1628

Qy       1602 HTGPWKPCTAACGRGFQSRKVDCT                HIRSCKPVAKRHC 1638
Db       1629 RVGNWGSCSVSCGVGVMQRSVQCLINEDQPSHLCHT DLKPEERKTC 1674


RESULT 3
US-09-930-872-4
; Sequence 4, Application US/09930872
; Patent No. 6448388
; GENERAL INFORMATION:
;  APPLICANT: Friddle, Carl Johan
;  APPLICANT: Hilbun, Erin
;  TITLE OF INVENTION: Novel Human Proteases and Polynucleotides Encoding the Sa
;  FILE REFERENCE: LEX-0219-USA
;  CURRENT APPLICATION NUMBER: US/09/930,872
;  CURRENT FILING DATE: 2001-08-14
;  PRIOR APPLICATION NUMBER: US 60/225,852
;  PRIOR FILING DATE: 2000-08-16
;  NUMBER OF SEQ ID NOS: 5
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
;   LENGTH: 1224
;   TYPE: PRT
;   ORGANISM: homo sapiens
US-09-930-872-4

  Query Match             8.7%;  Score 812.5;  DB 4;  Length 1224;
  Best Local Similarity   29.1%;  Pred. No. 5.2e-53;
  Matches  225;  Conservative   87;  Mismatches  276;  Indels  185;  Gaps   31;

Qy         59 PYDEQISPNIRGGEEKIGNWLAWGDWSECSRTCGGGASYSLRRCLT        GRNCEGQN 112
Db        576 KYGDEGPKPIH       GHWSDWSSWSPCSRTCGGGVSHRSRLCTNPKPSHGGKPCEGST 629

Qy        113 TRYKTCNHECPTAPEFRALQCTAYNEVQYQCHYYEWLPFYNGTAAPCALKCHALGQN 171
Db        630 RTLKTCNSQKCPRGLVEFRAAQCAEHNGRETRGEHYKWKPYTQVEDQDLCKLYCRAEGFD 689

Qy        172 LVVELAPKVLDGTRCNTDSLDMCISGICQAVGCDRQLGSNAKEDNCGVCAGDGSTCRLVR 231
Db        690 FFFSLSNKVKDGTPCSEDSRNVCIDGICERVGCDNVLGSDAVEDVCGVCNGNNSACTIHR 749

Qy        232 GQSKSHVSPEKREENVIAVPLGSRSVRITVKGPAHLFIESKTLQGSKGEHSFNSPGV 288
Db        750 GLYTKHHHHTNQYYHMVTIPSGARSIRIYEMNVSTSYISVRNALRRYYLNGHWTVDWPGR 809

Qy        289 FVVENTTVEFQRGSERQTFKIPGPLMADFIFKTR      YTAAKDSVVQFFFYQP 336
Db        809 YKFSGTTFDYRRSYNEPENLIAPGPTNETLIVELLFQGRNPGVAWEYSMPRLGTEKQPPA 868

Qy        337 QPISHQWEQTDTTECTVTCGQGYQLRNGADTVLTRLKEVVYPGHYCHYYPENVKPKPKPKPC 396
Db        869 QP SYTWAIVR SECSVSCGGG QMTVREGCYRDLKFQVNMSFCN    PKTRPVTGLVPC 922

Qy        397 SMDPCPSSDGFKEIMPYDHFQPLPRWEHNPWTACSVSCGGGIQRRSFVCVEESMHGEILQ 456
Db        923 KVSACP                     PSWSVGNWSACSRTCGGGAQSRPVQCTRR VHYDSE 964

Qy        457 VEEWKCMYAPKPKVMQTCNLFDCPKWIAMEWSQCTVTCGRGLRYRVVLCINHRGEH 511
Db        965 VPASLCPQPAPSSRQACNSQSCPPAWSAGPWAECSHTCGKGWRKRAVACKSTNPSARAQ 1023

Qy        512 HV  GGCNPQLKPHIKEECVIPIPCYKPKEKSPVEAKLPWLKQAQELEETRIATEEPTF 569
Db       1024 LLPDAVCTSEPKPRMHEACLLQ RCHKPK         KLQWLVSA              1059
```

```
Qy        570 PEPWSACSTTCGPGVQVREVKCRVLLTFTQTETELPEEECE    GPKLPTERPCLLEACD 626
Db       1060     WSQCSVTCERGTQKRFLK ALKYV SGKYEELASKKCSHLPKPSLELRAC      1109

Qy        627 ESPASRELDIPLPEDSETTYDWEYAGFTPCTATCLGGHQEATAVCLHIQTQQTVNDSLCD 686
Db       1110          APLP                                          CP 1115

Qy        687 MVHRPPAMSQACNTEPCPPRWHVGSWGPCSATCGVGIQTRDVYCLHPGETPAPPEECRD 745
Db       1116 -   RHPPFAAA    GPSRCSWFASEWSQCTASCGGGVQTRSVQCLAGGR    PASGCLLH 1166

Qy        746 EKPHALQACNQFDCPPG              W              HTEEWQQCSRTC 775
Db       1167 QKPSASLACNTHFCPIAEKKDAFCKDYFHWCYLVPQHGMCSHKFYGKQCCKTC 1219


RESULT 4
US-09-369-364A-17
; Sequence 17, Application US/09369364A
; Patent No. 6391610
; GENERAL INFORMATION:
;  APPLICANT: Apte, Suneel
;  APPLICANT: Hurskainen, Tiina L.
;  APPLICANT: Hirohata, Satoshi
;  TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases
;  FILE REFERENCE: 26473/4007/10-30-00
;  CURRENT APPLICATION NUMBER: US/09/369,364A
;  CURRENT FILING DATE: 1999-08-06
;  NUMBER OF SEQ ID NOS: 31
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
;   LENGTH: 1081
;   TYPE: PRT
;   ORGANISM: Homo sapiens ADAMTS-10
US-09-369-364A-17

  Query Match             8.5%;  Score 798.5;  DB 4;  Length 1081;
  Best Local Similarity   29.8%;  Pred. No. 5e-52;
  Matches  193;  Conservative   90;  Mismatches  221;  Indels  143;  Gaps   24;

Qy         72 EGKDGNWLAWGDWSECSRTCGGGASYSLRRC        LTGRNCEGQNTRYKTCSNHDCPP 125
Db        521 EGVDGAWGPWGPWGDCSRTCGGGVQFSSRDCTRPVPRNGGKYCEGRRTRFRSCNTEDCPP 580

Qy        126 DALDFRADQCTAYNEVQYQCHYYEWLPFYNGTAAPCALKCHAQGQNLVVELAPKVLDGTR 185
Db        581 GSQDFREVQCAEFDGAEVQGRRFYKWKPYTGGGVKACSLTCLAEGFNFYTERAAAVVDGTP 640

Qy        186 CNTDSLDMCISGQCQAVGCDRQLGSNAKEDNCGVCAGDGSTCRLVRGQSKSHVSPEKREE 245
Db        641 [illegible]                                                  [illegible]

Qy        246 NVIAVPLGSRSVRITVKGPAHLFI                       ESKTLQGSKGEHSFNSPGVF 289
Db        703 DVVWIP                KGSVHIFIQDLNLSLSHLALKGDQESLLLEGLPGT    PQPHRL 746

Qy        290 VVENTTVEFQRGSEE QTFKIPGPLMADFIF     KTRYTAAKDSVVQFFFYQPIS     340
Db        747 PLAGTTFQLRQGPDQVQSLEALGPINASLIVMVLARTELPA       LRYRFNAPIARDSL 801

Qy        341     HQWRQTDFTPCTVTCGGGYQLNSAETVLTRLKRVVYPGHYCHYYPENVKPKPKLKECSM 398
Db        802 PPYSWHYAPWTKCSAQCAGGSQVQAVECRNQLDSSAVAPHYCSAHSK   TTKRQRACNT 858

Qy        399 [illegible]                                    QRRSFVCVEESMHGEILQVE 458
Db        859 EPCP-  -  -  -  -       -  PDWVVGNWSLCSRSCDAGVRSRSVVCQRRVSAAEEKALD 904

Qy        459 EWKCMYAPKPKVMQTCNLFDCPKWIAMEWSQCTVTCGRGLRYRVVLCINHRGEHVGGCN 517
Db        905 DSACPQ QPRPPVLEACHGPTCPPEWATLDWSECTPSCGPGLRHRVVLCKS          956
```

```
Qy        518 PCLFIHIFFFTVII      IP YEPEEFSPVEALEPWEKQA QITFFIEIATEEPIFIPEPW 573
                        :   ::      :|  ||    | :   |        :       |  |  |         |
Db        951 ------ADQRSIIPPCHCIPAAKP--PSTMPCNI        PFCPPARWVISE-------W 989

Qy        574 SACSITCGPGVQVPEVECEVIITPTQTPTPIFFFFFTPSPFIPTE  PPCLLEACLESPAS 631
                 ||| || |  | | |:|            : |  |  ||        ||  ||  :: |:   |
Db        990 GFCSTQCGIGQQQRTVRC-------TSHTGQPSRPC------TFAIRPSIMQQCR---AK 1033

Qy        632 PELPIPLPPDSPTIYIWPYACPIP                ---CIATCLC 662
                 :  :|  :  |    |      : |              |  || |
Db       1034 CDSVVPPGDGPEECPIVNKVAYCPLVLEPLPCSRAYFRQMCCKTTQC 1080
```

RESULT 5
US-09-491-522-11
; Sequence 11, Application US/09491522
; Patent No. 6428998
; GENERAL INFORMATION:
;   APPLICANT: Colige, Alain
;   APPLICANT: Lapiere, Charles M.
;   APPLICANT: Prockop, Darwin J.
;   TITLE OF INVENTION: RECOMBINANT N PROTEINASE,
;   TITLE OF INVENTION: AND THE PRODUCTION, METHODS AND USES THEREOF
;   NUMBER OF SEQUENCES: 17
;   CORRESPONDENCE ADDRESS
;   ADDRESSEE: Pennie & Edmonds, LLP
;   STREET: 1155 Avenue of the Americas
;   CITY: New York
;   STATE: NY
;   COUNTRY: USA
;   ZIP: 10036-2811
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Diskette
;   COMPUTER: IBM Compatible
;   OPERATING SYSTEM: Windows
;   SOFTWARE: FastSEQ for Windows Version 2.0b
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/491,522
;   FILING DATE:
;   CLASSIFICATION:
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 08/886,333
;   FILING DATE:
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Abrams, Samuel B
;   REGISTRATION NUMBER: 30,605
;   REFERENCE/DOCKET NUMBER: 8389-0060-999
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 650-493-4935
;   TELEFAX: 650-493-5556
;   TELEX: 66141 PENNIE
;   INFORMATION FOR SEQ ID NO: 11:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 1205 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
US-09-491-522-11

```
  Query Match              6.4%;  Score 595.5;  DB 4;  Length 1205;
  Best Local Similarity   27.6%;  Pred. No. 1.9e-36;
  Matches  176;  Conservative   88;  Mismatches  267;  Indels  107;  Gaps   26;

Qy         74 KDGNWDAWCDWSDCSRTCGGGASYSLRRCLT------GRNCEGQNIRYKTCSNHDCPPDA 127
                 :||||  ||   :   ||||||  |    :   |:|          ||  |  |     :: |:: |||
Db        554 RDGNWGAWSPFGSCSRTCGIGVKFRTRQCDNPHPANGGRTCSGLAYDFQLCNSQDCPDAL 613

Qy        128 EDFRAQQCSAYN        DVQYQGHYYEWLP RYNDPAAPCALKCHAQGQNLVVELAPKV 180
                 ||| :||   ::         | |        : |||   : |      | |  |  | ::      || :     |
Db        614 ADFREEQCRQWDLYFEHGDAQ-----HHWLPHEHRDAKERCHLYCESKETGEVVSMKRMV 668

Qy        181 LDGTRCN TDSLDMCISGICQAVGCDRQIGSNAKEDNCGVCAGDGSTCRLVRGQSKSHVS 239
                 |||||:   |:   :|:  |  |:  ||||    :||:  :||  ||||  ||  |  |:::|:|  :  |
Db        669 HLGIRCSYKIAPSLVKGQCRKVCQGVICSKRQLRKCGVCGGENSHCKVVKGTFSRSP 727

Qy        240 PEFFFLNVIAVPLGSRSVRIIVKGPAHLFIESKTLQGSK   GEHSFNCPG   VFVVENT 294
                 :       :  :| ||:| : |                   :  | |:   |      |  :      |     |:
Db        728 KKLGYIKMFEIPAGARHLLIQEAEPSSHHLAVKNLLTGKFILNEENEVLGNSKIFIAMGV 787

Qy        295 TVEPQRCSERQTFKIPCPLM                     APPIPKTRYTAAKDSV      VQPFFYQP 338
                 |::       |:|  :   |||                      |          :|      :||:
Db        788 EWEYRLEEGRETLQTMGPLHCTFTVLVIPENDAKITLTYKYMIHEDGLNVLGNNVLEHGS 847

Qy        339 LSHQWKQTGFTECIVTCGGGYQLNSAHCVLIHLERVVTHRYCHYYLENVKEKEEEHEHG 397
                 :  ::|       :  ||:    ||||  |         |          ::|     :|       ::|  |||     : |:
Db        848 VGYEWALFEWSPCSKPCGGGSQFTKYGCRRRLDHKMVHRGFC----DSVSKPKAIRRTCN 903

Qy        398 MDFCPSSLQGFREIMQTGHFGLLRWLHNGWIACSVSCGGGIQRKSFVCVEESMHGEILQ 456
                     |                            |  |        |    ||  |||    |:|  ||    ||:        :
Db        904 PQECSQ-----------------PVWVTGEWEPCSRSCGRTGMQVRSVRCVQPLHNNTTRS 947

Qy        457 VEEWKCMYAPKPKVMQTCNLFDCP KWIAMEWSQCTVTCGRGLRYRVVLCINHRGEHVGG 515
                 |       |      |  :|:     :  ||        ||  :|  |    ||||:||||  |  :  |  |||             :    |
Db        948 VHTKHCNDA-RPPGRRACNRELCPGRWRAGSWSQCSVTCGNGTQERPVLC RTADDSFGV 1005

Qy        516 CNPQLKIHIKPECVIPIPCYK      PKEKSPVEAKLPWLKQAQE---LEET----RIATE 564
                 |   :   :               |  :    ||::          |  :||  |         :  ||   :            ::||             |        :
Db       1006 CREE-RPETARTCRLG PCPKNISDPSKKSYV - VQWLSRPDPNSPVQEISSKGRCQGD 1060

Qy        565 EPTP-         - -     IPEPWSACSITCGPGVQVREVKCRV -               LLTPTQTET 602
                 :   |                            ||          |     :|   |      :   :|    |                                |       |   :
Db       1061 KSVFCRMEVLSRYCSIPGYNKLCCKSCNPHDNLTDVDDRAPPSCKHNDIEELMPTLSVP 1120

Qy        603 ELPEEECEGPKLPTERP    CLLEACDESPASRELDIP 637
                  |    |          |  :|  |  |                  |  ::   |   :       :|:|
Db       1121 TLVMEVQPPPGIPLEVPLNTSSTNATEDHPEINAVDVP 1158
```

RESULT 6
US-09-130-491-2
; Sequence 2, Application US/09130491
; Patent No. 6416974
; GENERAL INFORMATION:
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Goodearl, Andrew D.J.
; TITLE OF INVENTION: TANGO-71, TANGO-73, TANGO-74, TANGO-76, AND TANGO-83
; FILE REFERENCE: 09404/041001
; CURRENT APPLICATION NUMBER: US/09/130,491
; CURRENT FILING DATE: 1998-08-07
; EARLIER APPLICATION NUMBER: US 60/058,108
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: US 60/054,961
; EARLIER FILING DATE: 1997-08-06
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
;   LENGTH: 967
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-130-491-2

```
  Query Match              6.1%;  Score 571.5;  DB 4;  Length 967;
  Best Local Similarity   28.2%;  Pred. No. 8.8e-35;
  Matches  136;  Conservative   77;  Mismatches  164;  Indels  105;  Gaps   19;

Qy         76 GNWDAWCDWSDCSRTCGGGASYSLRRCLT          GRNCEGQNIRYKTCSNHDCPP DAE 128
                 |:|   || |  |||||||||    |::| |                 |:  |||:  :||::|:   |||    :  :
Db        560 GSWGMWGPWGDCSRTCGGGVQYTMRECDNPVPKNGGKYCEGKRVRYRSCNLEDCPDNNGK 619

Qy        129 DFRAQQCSAYNDVQY       QGHYYEWLPRYN   DPAAPCALKCHAQGQNLVVELAPKVLD 182
                  ||  :||  |:|:                 |        ||:|:|            |      |   |   |   |:|                 |   |||:|
Db        620 TFREEQCEAHNEFSKASFGSGPAVEWIPKYAGVSPKDRCKLICQAKGIGYFFVLQPKVVD 679

Qy        183 GTRCNTDSLDMCISGICQAVGCDRQIGSNAKEDNCGVCAGDGSTCRLVRGQSKSHVSPEK 242
                 ||  |:  ||    :|:  |  |      ||||  :  |     |  |  ||||   |:||||:  :  |      |    |   :
```

```
Db        706 GIPCSPDSTSVCVQGQCVKAGCDRTIDSKKKFDKCGVCGGNGSTCKKISG     SVTSAKP 756

Qy        243 REENVIAVPLGSRSVRITVKGPAHLFIESKTLQGSKGEHSF       NSPGVFVVEN     T 294
                  : : :|  :| |: : :           : :  :  :||:    ||      :    : : :     :
Db        757 GYHPTTTTPTGATNTF                VKQRNQRGGRNNGSTLATKAAGGYTLNGDYTLS 786

Qy        295 TVEFQRGSERQTFKIPGPLMADFIFK      TRYTAAKDSVVQFFFYQPISHQWRQTDFFPC 351
              |:|                           |  ::|      ||: :  :: :    : |:           |
Db        787 TLE                         QGIMYKGVVLRYSGSSAALERIRSFSPLKE          PL 821

Qy        352 TVTCGGGYQLNSADCVDIRLKRVVPCHYCHYYPENVKPKPKLKECSMEPCPSGDGFKFIM 411
              |:          |       :   : : .          | : :             | : :  '  :   |:
Db        822 TIQV        TIVGNALRPKIK          YTYFV             KKKKESFNAIPT             854

Qy        412 PYDHPQPLPRWEHNPWTACSVSCGGGIQRRSFVCVEESMHGEILQVEEWKCMYAPKPKVM 471
                          |       |   || ||   | ||     |        ::|:          :|      ||
Db        855             FSAWVIEEWGECSKSCELGWQRRLVEC   RDINGQPAS      ECAKEVKPAST 901

Qy        472 QTCNLFDCPKWIAMEWSQCTVTCGRGLRYRVVLCINHRGEHVG   GCNPQLK LHIKEEC 528
              : |       ||:|      ||' :  : |'|:|  : |  :  :::| |   :       |:|   |         :
Db        902 RPCADHPCPQWQLGEWSSCSKTCGKGYKKRSLKCLSHDGGVLSHESCDPLKKPKHFIDFC 961

Qy        529 VI 530
               .
Db        962 TM 963


RESULT 7
US-09-130-491-16
; Sequence 16, Application US/09130491
; Patent No. 6416974
; GENERAL INFORMATION:
;  APPLICANT: Holtzman, Douglas A.
;  APPLICANT: Goodearl, Andrew D.J.
;  TITLE OF INVENTION: TANGO 71, TANGO 73, TANGO 74, TANGO 76, AND TANGO 83
;  FILE REFERENCE: 09404/041001
;  CURRENT APPLICATION NUMBER: US/09/130,491
;  CURRENT FILING DATE: 1998-08-07
;  EARLIER APPLICATION NUMBER: US 60/058,108
;  EARLIER FILING DATE: 1997-09-05
;  EARLIER APPLICATION NUMBER: US 60/054,961
;  EARLIER FILING DATE: 1997-08-06
;  NUMBER OF SEQ ID NOS: 16
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 16
;   LENGTH: 551
;   TYPE: PRT
;   ORGANISM: Rattus rattus
US-09-130-491-16

  Query Match             6.0%;  Score 564.5;  DB 4;  Length 551;
  Best Local Similarity   27.7%;  Pred. No. 1.3e-34;
  Matches  133;  Conservative   72;  Mismatches  173;  Indels  103;  Gaps   18;

Qy         76 GNWDAWGDWSDCSRTCGGGASYSLRRCLT--    GRNCEGQNIRYKTCSNHDCPPDAE 128
              |:.   || |   |:|||||| |::| |             |: |||: :||:::: |||  : :
Db        144 GSWGPWGPWGDCSRTCGGGVQYTMRECDNPVPKNGGKYCEGKRVRYRSCNIEDCPDNNGK 203

Qy        129 DFRAQQCSAYNDVQYQGH      YYEWLPRYN   DPAAPCALKCHAQGQNLVVELAPKVLD 182
               || :|| |:|:                ||  |:|       |    | |  : |:|        | ||.:|
Db        204 TFREEQCEAHNEFSKASFGNEPTVEWTPKYAGVSPKDRCKLTCEAKGIGYFFVLQPKVVD 263

Qy        183 GIPCNIDSLDMCISGICLAVGCDRQLGSNAKEDNCGVCAGDGSTCRLVRGQSKSHVSPEK 242
              |: :   ||   :|: | |     |:||  :  |   | | |:|| |: :|:|:  :  :
Db        264 GIPCSPDSTSVCVQGQCVKAGCDRIIDSKKKFDKCGVCGGNGSTCKKMSG     TVTSTRP 320

Qy        243 KEENVIAVPLGSRSVRITVKGPAHLFIESKTLQGSKGEHSF       NSPGVFVVENTTVEF 298
                  : : :|  :| |: :          : :  :  :||:    ||      :  |  : : :        |
Db        321 GYHDIVTIPAGATNTF               VKHRNQRGSRNNGSTLATKAAGGYTLNG     NF 367

Qy        299 QRGSERQTFKIPGPLMADFIFKTRYTAAKDSVVQFFFYQPISHQWRQTDFFPCTV TCGG 357
                 :  |       | : :             ||: :  : :  :  : : |:                | |:
Db        368 TLSTLEQDLIYKGIVL          RYSGSSAALERIRSFSPLKE          PLTIQVLMV 412
```

```
Qy        358 GYQLNSADCVDIRLKRVVPTTYCHYYPENVKPKPKLKECSMEPCPSGDGFKLIM   PYDH 415
              |: |                                  :||:|                  |    |    :
Db        413 GHAL                                  PPKIK                TIYPMKKKIES 432

Qy        416 PQPLP      RWEHNPWTACSVSCGGGIQRRSFVCVEESMHGEILQVEEWKCMYAPKPKVMQ 472
              |   :|       |      |   || :||  | |||   .  :  :  :                :|      ||   :
Db        433 FNAIPTFSEWVIEEWGECSKTCGSGWQRRVVQCRDINGH        PASECAKEVKPASTR 486

Qy        473 TCNLFDCPKWIAMEWSQCTVTCGRGLRYRVVLCINHRGEHVG   GCNPQLK LHIKEECV 529
                |         || |     :||  |:  |||:|  :  :  : |::|  |     :      |:|   |    |   : |
Db        487 PCADLPCPHWQVGDWSPCSKTCGKGYKKRTLKCVSHDGGVLSNESCDPLKKPKHYIDFCT 546

Qy        530 I 530
              :
Db        547 L 547


RESULT 8
US-09-491-522-5
; Sequence 5, Application US/09491522
; Patent No. 6428998
;  GENERAL INFORMATION:
;     APPLICANT:  Colige, Alain
;     APPLICANT:  Lapiere, Charles M.
;     APPLICANT:  Prockop, Darwin J.
;     TITLE OF INVENTION:  RECOMBINANT N-PROTEINASE,
;     TITLE OF INVENTION:  AND THE PRODUCTION, METHODS AND USES THEREOF
;     NUMBER OF SEQUENCES:  17
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Pennie & Edmonds, LLP
;       STREET:  1155 Avenue of the Americas
;       CITY:  New York
;       STATE:  NY
;       COUNTRY:  USA
;       ZIP:  10036-2811
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Diskette
;       COMPUTER:  IBM Compatible
;       OPERATING SYSTEM:  Windows
;       SOFTWARE:  FastSEQ for Windows Version 2.0b
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/491,522
;       FILING DATE:
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/886,333
;       FILING DATE:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Abrams, Samuel B
;       REGISTRATION NUMBER:  30,605
;       REFERENCE/DOCKET NUMBER:  8389-0060-999
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  650-493-4935
;       TELEFAX:  650-493-5556
;       TELEX:  66141 PENNIE
;   INFORMATION FOR SEQ ID NO:  5:
;     SEQUENCE CHARACTERISTICS
;       LENGTH:  1211 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
US-09-491-522-5

  Query Match             5.9%;  Score 553.5;  DB 4;  Length 1211;
  Best Local Similarity   28.9%;  Pred. No. 2.5e-33;
  Matches  153;  Conservative   71;  Mismatches  213;  Indels   93;  Gaps   22;

Qy         74 KGGNWGAWGDWSDCSRTCGGGASYSLRRCLT        GRNCEGQNIRYKTCSNHDCPPDA 127
                :||:| ||   :   |||||. |   :   ::|             .| | |        :  :  .  |||
Db        560 RDGSWGAWSPWSSCSRTCGGGVKFRTRQCDNPHPANGGRTCSGLAYDFQLCSRQDCPDSL 619

Qy        128 EDFRAQQCSAYN         DVQYQGHYYEWLP  RYNDPAAPCALKCHAQGQNLVVELAPKV 180
```

```
             ||| :||  ::          | |       : |||  : |     | | | ::     || :   |
Db      620 AGFREDQCRQWDLYFEHGDAQ      HHWLPHEHRDAKERCHLYCESRETGEVVSMKRMV 674

Qy      181 LDGTRCN TLSLDMCISGICQAVGCDRQLGSNAKEDNCGVCAGDGSTCRLVRGQSKSHVS 239
             :||||:  |:  : :  | | :  ||||  :||: :.| |||| || | |::|:|         |
Db      675 HDGTRCSYKDAFSLCVRGDCRKVGCDGVIGSSKQEDKCGVCGGDNSHCKVVKGTFTR  S 732

Qy      240 PEKREE NVIAVPLGSRSVRITVKGPAHLFIESKTLQGSK----GEHSFN-SPGVFVVEN 293
             |:|       :   :|   :| : |               :  | |: |         |:  : |   |:
Db      733 PKKHGYIKMFEIPAGARHLLIQEVDATSHHLAVKNLETGKFILNEENDVDASSKTFIAMG 792

Qy      294 TTVEFQRGSERQIFKIPGPLMADF------IFKTRYTAA----------------  KD 328
               |::     |:|  :   |||                :   || :                  :|
Db      793 VEWEYRDEDGRETLQTMGPLHGTITVLVIPVGDTRVSLTYKYMIHEDSLNVDDNNVLEED 852

Qy      329 SVVQFFFYQPISHQWRQTDFFPCTVTCGGGYQLNSAECVDIRLKRVVPDHYCHYYPENVK 388
             |||              ::|     : ||:  |||| |        |           ::|    :|          |
Db      853 SVV---------YEWALKKWSPCSKPCGGGSQFTKYGCRRRLDHKMVHRGFCAALS---K 900

Qy      389 PKPKLKECSMDPCPSSDGFKEIMPYDHFQPLPRWEHNPWIACSVSCG GGIQRRSFVCVE 447
             ||    : |: | |                                  | |      |  || :||   |:| ||   |::
Db      901 PKAIRRACNPDECSQ-----------------PVWVTGEWEPCSQTCGRTGMQVRSVRCIQ 944

Qy      448 ESMHGEILQVEEWKCMYAPKPKVMQTCNLFDCP-KWIAMEWSQCTVTCGRGLRYRVVLCI 506
                         |      ,   |  :|:   : |:      || :| |   ||||:|||| |  :  | |||
Db      945 PLHDNTTRSVHAKHCNDA RPESRRACSRELCPGRWRAGPWSQCSVTCGNGTQERPVLC- 1002

Qy      507 NHHGEHVGGCNPQLKLHIKEECVIPIPCYK---PKEKSPVEAKLPWLKQ 552
                   :   | |   :  :         | :    || :         | :|' |      :  || :
Db     1003 RTADDSFGICQEE-RPETARTCRLG-PCPRNISDPSKKSYV   VQWLSR 1047
```

```
RESULT 9
US-09-369-364A-19
; Sequence 19, Application US/09369364A
; Patent No. 6391610
; GENERAL INFORMATION:
;  APPLICANT:  Apte, Suneel
;  APPLICANT:  Hurskainen, Tiina L.
;  APPLICANT:  Hirohata, Satoshi
;  TITLE OF INVENTION:  Nucleic Acids Encoding Zinc Metalloproteases
;  FILE REFERENCE:  26473/4097/10-30-00
;  CURRENT APPLICATION NUMBER: US/09/369,364A
;  CURRENT FILING DATE:  1999-08-06
;  NUMBER OF SEQ ID NOS:  31
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 19
;   LENGTH: 450
;   TYPE: PRT
;   ORGANISM: Mus musculus ADAMTS-19
;   FEATURE:
;   NAME/KEY: MOD_RES
;   LOCATION: (113)
;   OTHER INFORMATION: Xaa   H
;   NAME/KEY: MOD_RES
;   LOCATION: (118)
;   OTHER INFORMATION: Xaa   A
US-09-369-364A-19

  Query Match             5.8%;  Score 545;  DB 4;  Length 450;
  Best Local Similarity   24.4%;  Pred. No. 2.3e-33;
  Matches  162;  Conservative   66;  Mismatches  185;  Indels  252;  Gaps   25;

Qy      177 APKVLDGTRCNTDSLDMCISGICQAVGCDRQLGSNAKEDNCGVCAGDGSTCRLVRGQSKS 236
             |  |:||||  |   |::|:|:|| |: ||||| |||: :|| |  || |||| |   :  |    |
Db        1 AAAVVDGTPCRPDTVDTCVSGECKHVGCDRLLGSDLREDKCRVCGGDGSACETIEGVF-S 59

Qy      237 HVSPEKREENVIAVPLGSRSVRITVKGPAHLFI----------------ESKTLQGSKGE 280
               |     |:|: :|              ||  |:||                   ||  |:|   |
Db       60 PALPGTGYEDVVWIP-----------KGSVHIFIQDLNLSLSHLALKGDQESLLLEGLPGT 109

Qy      281 HSFNSPGVFVVENTTVEFQRGSERQIFKIPGPLMADFIFKTRYTAAKDSVVQFFFYQPI 339
                 |     :  ||    ::| :: |: :   ||: |  |            |:    : :  |  ||
```

```
Db      110 ---PQPXRLPLXGTTFHLRQGPDQAQSLEALGPINASLTIMV LAQAELPALHYRFNAPI 165

Qy      340 S------HQWRQTDFFPCTVTCGGGYQLNSAECVDIRLKRVVPDHYCHYYPENVKPKPKL 393
             :         : |      :    |:   | || |:      || :             |   |||   : :      || :
Db      166 ARDALPPYSWHYAPWTKCSAQCAGGSQVQVVECRNQLDSSAVAPHYCSGHSK  LPK RQ 222

Qy      394 KECSMDPCPSSDGFKEIMPYDHFQPLPRWEHNPWTACSVSCGGGIQRRGFVCVEESMHGE 453
             : |: :|||                                   | |      |: || ||   |:: || ||            |
Db      223 RACNTEPCP- ---                  PDWVVGNWSRCSRSCDAGVRSRSVVCQRRVSAAE 265

Qy      454 ILQVEEWKCMYAPKPKVMQTCNLFDC PKWIAMEWSQCTVTCGRGLRYRVVLCINHHGEH 512
                  :::    |     |:| |:: |        | |:|    ::||:|| :|| |||:|||||
Db      266 EKALDDSACP-QPRPPVLEACQGPMCPPEWATLDWSECTPSCGPGLRHRVVLC---  -- 317

Qy      513 VGGCNPQLKLHIKEECVIPIPCYKPKEKSPVEAKLPWLKQAQELEETRIATEEPTFIPEP 572
                                                                  | | :
Db      318 ------------------------------------KSADQ   -    -        -  322

Qy      573 WSACSTTCGPGVQVREVKCRVLLTFTQIETELPEEECEGPKLPTERPCLLEACDESPASR 632
                                               : ||     |        ||    :|
Db      323 -------------------------RSTLPPGHC----LPAAKP------ --     -- 337

Qy      633 ELDIPLPEDSETTYDWEYAGFTPCTATCLGGHQEAIAVCLHIQTQQTVNDSLCDMVHRPP 692
                                         | |
Db      338 ----------------------PST-    -   --------- ------- -------    --- 340

Qy      693 AMSQACNTEPCPP-RWHVGSWGPCSATCGVGIQTRDVYCL HPGETPAPPEECRDE KPH 749
                   ||     ||| ||       || ||   ||:| | | | |    | |:       |   || :   :|
Db      341 ---MRCNLRRCPPARWVTSEWGECSTQCGLGQQQRTVRCTSHTGQ-- PSRECTEALRPS 394

Qy      750 ALQACNQFDC      PPGWHIEEWQQCSRTCGGGTQNRRVTCR        QLLTDGSFLNLSD 800
              :| |  :   |      |||    || :                   |:      |          |   :      |    :|
Db      395 TMQQC-EAKCDSVVPPGDGPEECKDV-   -       NKVAYCPLVLKFQFCSRAYFRQMSC 444

Qy      801 ELCQG 805
             : |||
Db      445 KTCQG 449
```

```
RESULT 10
US-09-369-364A-5
; Sequence 5, Application US/09369364A
; Patent No. 6391610
; GENERAL INFORMATION:
;  APPLICANT:  Apte, Suneel
;  APPLICANT:  Hurskainen, Tiina L.
;  APPLICANT:  Hirohata, Satoshi
;  TITLE OF INVENTION:  Nucleic Acids Encoding Zinc Metalloproteases
;  FILE REFERENCE:  26473/4097/10-30-00
;  CURRENT APPLICATION NUMBER: US/09/369,364A
;  CURRENT FILING DATE:  1999-08-06
;  NUMBER OF SEQ ID NOS:  31
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 5
;   LENGTH: 859
;   TYPE: PRT
;   ORGANISM: Homo sapiens ADAMTS-6
;   FEATURE:
;   NAME/KEY: MOD_RES
;   LOCATION: (450)
;   OTHER INFORMATION: Xaa = L
US-09-369-364A-5

  Query Match             5.5%;  Score 515;  DB 4;  Length 859;
  Best Local Similarity   34.9%;  Pred. No. 1.5e-30;
  Matches  109;  Conservative   46;  Mismatches  119;  Indels   38;  Gaps    8;

Qy       75 DGNWDAWSPWSECSRTCGGGASYSLRRCL- -     LTPRNGGQNLRYKEFSNHDCPPDAE 128
             || |   |   | :|||||||| | |||| |              : | | |:  ||::|:    ||   :
Db      510 DGGWGPWSLWGECSRTCGGGVSSSLRHCDSPAPSEVEKYCLGERKRYRSCNTDPCPLGSR 569

Qy      129 DFRAQQCSAYNDVQYQGHYYEWLPRYNDPAAPCALKCHAQGQNLVVELAPKVLDGTRCNT 188
             ||| :||: :::: ::| || | |          |||| | |:| |     | || |:|||:||
Db        570 DFPFKLYAPTFNMPFPCKYYNWKPYIGGVKFCALNCLAEGYNFYTERAPAVIDGTQCNA 629

Qy        189 DSLDMCISGICQAVGCDRQLGSNAKEDNCGVCAGDGSTCRLVRGQSKSHVSPEKREENVI 248
              ||||:|| |  |: |||    ||::||| | ||  | |||| :  |      : |     |:
Db        630 DSLDICINGECKHVGCDNILGSDAREDRCRVCGGDGSTCDAIEGFFNDSL PRGGYMEVV 688

Qy        249 AVPLGSRSVRITVKGPAHLFIESKTLQGS     KGEHSFNSPGVFVVENTTVEFQRGS E 303
               :| ||  : :      :  :|   | : :|      |  : :  |   |  |  .   ::| : :
Db        689 QIPRGSVHIEVREVAMSKNYIALKS EGDDYYINGAWTIDWPRKFDVAGTAFHYKRPTDE 747

Qy        304 RQTFKTPGPLMADFIFKT              RY       IAAKDSVVQFFFYQPISHQWRQT 346
               : :  :   ||    : |              ||       | : |: | |           |
Db        748 PESLEALGPTSENLIVMVLLQEQNLGIRYKFNVPITRTGSGDNEVGF          TWNHQ 799

Qy        347 DFFPCTVTCGGG 358
               :   |:  || ||
Db        800 PWSECSATCAGG 811


RESULT 11
US-09-369-364A-15
; Sequence 15, Application US/09369364A
; Patent No. 6391610
; GENERAL INFORMATION:
;  APPLICANT:  Apte, Suneel
;  APPLICANT:  Hurskainen, Tiina L.
;  APPLICANT:  Hirohata, Satoshi
;  TITLE OF INVENTION:  Nucleic Acids Encoding Zinc Metalloproteases
;  FILE REFERENCE:  [illegible]
;  CURRENT APPLICATION NUMBER:  US/09/369,364A
;  CURRENT FILING DATE:  1999-08-06
;  NUMBER OF SEQ ID NOS:  31
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 15
;   LENGTH: 874
;   TYPE: PRT
;   ORGANISM: Mus musculus ADAMTS-9
US-09-369-364A-15

  Query Match             5.4%;  Score 506.5;  DB 4;  Length 874;
  Best Local Similarity   20.1%;  Pred. No. 6.8e-30;
  Matches  149;  Conservative   70;  Mismatches  188;  Indels  335;  Gaps   22;

Qy         75 DGNWDAWGDWSDCSRTCGGGASYSLRRCLT       GRNCEGQNIRYKTCSNHDCPPDAE 128
              ||:   |   :   ||||||||     ::| |          |: |  |: :::|:|:     |
Db        427 DGSWGCWSHFSTCSRTCGGGIKTATRECNRPEPKNGGKYCVGRRMKFKSCNIEPCMKQKR 486

Qy        129 DFRAQQCSAYNDVQYQGHYY               EWLPRYNDPAAP  CALKCHAQGQNLVVELA 177
              ||| :|||:            : | ::              | |:|:          | | |   |     :|
Db        487 DFREEQCA      HFDGKHFNINGLLPSVRWIPKYSGILMKDRCKLFCRVAGNTAYYQLR 541

Qy        178 PKVLDGTRCNTDSLDMCISGICQAVGCDRQLGSNAKEDNCGVCAGDGSTCRLVRGQSKS  236
               :|:|||| |    : |:|: :|:   .|:  | |   ::| |||:| || |:|:  :    :
Db        542 DRVIDGTPCGQDTNDICVQGLCRQAGCDHILNSKVRRDKCGICGGDNSSCKTVAGTFNTV 601

Qy        237 HVSPEKREENVIAVPLGSRSVRITVKGPAHLFIESK         TLQGSKGEHSFNSPGVFV 290
              .            |: :| |: |: :      ||             |  :||:   |   : :
Db        602 HYG      YNIVVRIPAGATSIDVR      QHSFSGKSEDDNYLALSNSKGEFLLN  GDFV 651

Qy        291 VENTTVEFQRGSERQTFKTPGPLMADFIFKTRYTAAKDSVVQFFFYQPISHQWRQTDFFP 350
              |       | : ||                                :|:::
Db        652 VSMSKREVRVGS                                AVIEY          668

Qy        351 CTVTCGGGYQLNSAECVDIRLKRVVPDHYCHYYPENVKPKPKLKECSMDPCPSSDGFKEI 410
                              | :
Db        669      GGGDNVCERLNCTD                                          682

Qy        411 MPYDHFQPLPRWEHNPWTACSVSCGGGIQRRSFVCVEESMHGEILQVEEWKCMYAPKPKV 470
                                                      :||  :   ::| |
Db        683                                        RIEEELLLQVLSV          695

Qy        471 MQTCNLFDCPKWIAMEWSQCTVTCGRGLRYRVVLCINHRGEHVGGCNPQLKLHIKEECVI 530
                                  |   |                    || ::          |
```

```
Db        696                          GKLY                 NPDVRYSFN    I 709

Qy        531 PIPCYKPKEKSPVEAKLPWLKQAQELEETRIATEEPTFIPEPWSACSTTCGPGVQVREVK 590
               ||      |   |    :                              || |||
Db        710 PT       EDKP  QQFYWNSHG                        PWQACS       733

Qy        591 CRVLLTFTQTETELPEEECEGPKLPTERPCLLEACDESPASRELDIPLPEDSETTYDWEY 650
                                : :  : |
Db        734 ------           KPQQERRPK                                   740

Qy        651 AGFTPCTATCLGGHQEATAVCLHIQTQQTVNDSLCDMVHRPPAMSQACNTEPCPPRWHVG 710
                                  |        | ||:|  || : :|  ::||| |: |  ||||
Db        741 ------------    ---LVCTRESDQLTVSDQRCDRLPQPGPVTEACGTD CDLRWHVA 781

Qy        711 SWGPCSATCGVGIQTRDVYCLHPGETPAPPEECRDE- - ---KPHALQACNQFDCPPGWH 764
               |   ||| ||:: :| |:.           |:  |          :|   : |:          ||
Db        782 SKSECSAQCGLGYRTLDIHCAKYSRMDGKTEKVDDSFCSSQPKPSNQEKCSGECSTGGWR 841

Qy        765 IEEWQQCSRTCGGGTQNRRVTC 786
                  | :||||:| |||  :  |
Db        842 YSAWTECSRSCDGGTHRRRAIC 863


RESULT 12
US-09-130-491-8
; Sequence 8, Application US/09130491
; Patent No. 6416974
; GENERAL INFORMATION:
;  APPLICANT: Holtzman, Douglas A.
;  APPLICANT: Goodearl, Andrew D.J.
;  TITLE OF INVENTION: TANGO 71, TANGO 73, TANGO 74, TANGO 76, AND TANGO 83
;  FILE REFERENCE: 09404/04100
;  CURRENT APPLICATION NUMBER: US/09/130,491
;  CURRENT FILING DATE: 1998-08-07
;  EARLIER APPLICATION NUMBER: US 60/058,108
;  EARLIER FILING DATE: 1997-09-05
;  EARLIER APPLICATION NUMBER: US 60/054,961
;  EARLIER FILING DATE: 1997-08-06
;  NUMBER OF SEQ ID NOS: 16
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 8
;   LENGTH: 481
;   TYPE: PRT
;   ORGANISM: Rattus rattus
US-09-130-491-8

  Query Match             5.3%;  Score 499.5;  DB 4;  Length 481;
  Best Local Similarity   28.3%;  Pred. No. 9.3e-30;
  Matches  130;  Conservative   66;  Mismatches  145;  Indels  119;  Gaps   20;

Qy         34 FLPEEFALSPQGSFLEDTTGEQFLIYRYDDQTSRNTRS    -                  70
              | |:|    |  ! :||    : :   :|   |:     :|::
Db         47 FGPDERHCPNTS VEDTCVQLWCRHRDSDEPTCHIKNASLLWADGIPCGPGELCLDGSCV 105

Qy         71 ---DEDK     DGNWDAWGDWSDCSRTCGGGASYSLRRCLT       GENCEGQNIRYK 116
                   :   |       ||:|  || |  ||||||||  :| | |            |: | |: ::|:
Db        106 LREEVENPKAVVDGDWGPWGPWGQCSRTCGGGIQFSNRECDNPAPQNGGRFCQGERVKYQ 165

Qy        117 TCSNHDCPPDAEDFRAQQC     SAYNDVQYQGHYYEWLPRYN  DPAAPCALKCHAQGQN 171
              :|     :|||: :  || ||     :|||          |:: :|:|:|:     |    | | | |:|::
Db        166 SCKTEECPPNGKGFREQQCEKYNAYNHTDLDGNFLQWVPKYSGVSPRDRCKLFCRARGRS 225

Qy        172 LVVELAPKVLDGTRCNTDSLDMCISGICQAVGCDRQLGSNAKEDNCGVCAGDGSTCRLVR 231
                        ||:|||| |   :| :|: |  |     ||    :  |   | | |||| ! . : || |
Db        226 EFKVFEAKVIDGTLCGPETLAICVRGQCVKAGCDHVVNSPKKLDKCGVCGGKGTACRKVS 285

Qy        232 GQSKSHVSPEKREENVIAVPLGSRSVRITVKGPAHLFIESK      TLQGSKGEHSFN SP 286
              |   |                 ::: :| |   :   | |  :|   :::          |: : |:: | :
Db        286 G    SFTPFSYGYNDIVTIPAG  ATNIDVKQRSHPGVQNDGSYLALKTANGQYLLNGNL 340

Qy        287 GVFVVENT        TVEFQRGS--    ER QTFK  TPGPLMADFIF           KTRYT    324
               :  :|            |:      ||      || |:|: :| ||        :               : :||
Db        341 AISATEQDILMKGTTLKYSGSMATLERLQSFQALPEPLTVQLLTVSGEVFPPKVKYTFFV 400
```

```
Qy     325 --------------AAKDSVVQFFFYQPTSHQWRQTDFFPCTVTCGGGYQLNSAEC----- 366
                         |   :::|    |       :|    |:   |   |||||:|   : ||
Db     401 PNDTDFNVQCCKERASTNIIQSLFYA----EWVLGEWSECPSTCGGGWQRRTVECRDPSS 456

Qy     367 -----VDIRLKRVVPDHYCHYYPENVKPKPKLKECSMDPC 401
                    |  ||            ||: ||      |   ||
Db     457 QASDTCDEALK        PEDAKF      CGSQFC 479


RESULT 13
US-09-130-491-13
; Sequence 13, Application US/09130491
; Patent No. 6416974
; GENERAL INFORMATION:
;  APPLICANT: Holtzman, Douglas A.
;  APPLICANT: Goodearl, Andrew D.J.
;  TITLE OF INVENTION: TANGO-71, TANGO-73, TANGO-74, TANGO-76, AND TANGO-83
;  FILE REFERENCE: 09404/041001
;  CURRENT APPLICATION NUMBER: US/09/130,491
;  CURRENT FILING DATE: 1998-08-07
;  EARLIER APPLICATION NUMBER: US 60/058,108
;  EARLIER FILING DATE: 1997-09-05
;  EARLIER APPLICATION NUMBER: US 60/054,961
;  EARLIER FILING DATE: 1997-08-06
;  NUMBER OF SEQ ID NOS: 16
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 13
;   LENGTH: 608
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-130-491-13

  Query Match             5.3%;  Score 499.5;  DB 4;  Length 608;
  Best Local Similarity   28.0%;  Pred. No. 1.3e-29;
  Matches  120;  Conservative   70;  Mismatches  138;  Indels  101;  Gaps   18;

Qy      76 GNWDAWGDWSDCSRTCGGGASYSLRRCLT------GRNCEGQNIRYKTCSNHDCPP-DAE 128
           |:|   || | |||||||||   |::| |          |: |||. .||.:|:  |||   
Db     255 GSWGMWGPWGDCSRTCGGGVQYIMRECDNPVPKNGGKYCEGKRVRYRSCNLEDCPDNNGK 314

Qy     129 DFRAQQCSAYNDVQY----QGHYYEWLPRYN--DPAAPCALKCHAQGQNLVVELAPKVLD 182
            || :|| |:|:              |    ||:|:|    |   | | | |.|      | |||:|
Db     315 TFREEQCEAHNEFSKASFGSGPAVEWIPKYAGVSPKDRCKLICQAKGIGYFFVLQPKVVD 374

Qy     183 GTRCNTDSLDMCISGICQAVGCDRQLGSNAKEDNCGVCAGDGSTCRLVRGQSKSHVSPEK 242
           || |: ||    :|: | |    |||| . |   | | |||| |.||||. . |      |   | .
Db     375 GTPCSPDSTSVCVQGQCVKAGCDRIIDSKKKFDKCGVCGGNGSTCKKISG    SVISAKP 431

Qy     243 REENVIAVPLGSRSVRITVKGPAHLFIESKTLQGSKGEHSF----NSPGVFVVEN----T 294
                ::| :|:|: ::                  :: :  :||:    ||       : | :::     :
Db     432 CYHDIITIPIGATNIE               VKQRNQRGSRNNGSFLAIKAADGTYILNGDYTLS 481

Qy     295 TVEFQRGSERQTFKIPGPLMADFIFK---TPYTAAKDSVVQFFFYQPISHQWRQTDFFPC 351
           |:|                    | ::|      ||:  :   :: :     : |:          |
Db     482 TLE-------          QDIMYKGVVLRYSGSSAALERIRSFSPLKE        PL 516

Qy     352 TVTCGGGYQLNSAECVDIRLKRVVPDHYCHYYPENVKPKPKLKECSMDPCPSSDGFKEIM 411
           |:        |       :   ::|         | ::               | |: |  :   |:
Db     517 TIQV-----LTVGNALRPKIK------YTYFV--------KKKKESFNAIPT-------- 549

Qy     412 PYDHFQPLPRWEHNPWTACSVSCGGGIQRRSFVCVEESMHGEILQVEEWKCMYAPKPK-V 470
                     |      |    || :|| | ::||  |:   |   | :|   |     |      |||
Db     550         FSAWVIEEWGECSKTCGKGYKKRSLKCL  SHDGGVLSHE  SCDPLKKPKHF 598

Qy     471 MQTCNLFDC 479
           :  | :  :|
Db     599 IDFCTMAEC 607


RESULT 14
US-09-369-364A-9
; Sequence 9, Application US/09369364A
```

```
; Patent No. 6391610
; GENERAL INFORMATION:
;  APPLICANT:  Apte, Suneel
;  APPLICANT:  Hurskainen, Tiina L.
;  APPLICANT:  Hirohata, Satoshi
;  TITLE OF INVENTION:  Nucleic Acids Encoding Zinc Metalloproteases
;  FILE REFERENCE:  26473/4007/10-30-00
;  CURRENT APPLICATION NUMBER:  US/09/369,364A
;  CURRENT FILING DATE:  1999-08-06
;  NUMBER OF SEQ ID NOS:  31
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 9
;   LENGTH: 905
;   TYPE: PRT
;   ORGANISM: Mus musculus ADAMTS-8
US-09-369-364A-9

  Query Match             5.2%;  Score 484.5;  DB 4;  Length 905;
  Best Local Similarity   27.9%;  Pred. No. 3.4e-28;
  Matches  128;  Conservative   54;  Mismatches  153;  Indels  113;  Gaps   19;

Qy      34 FLPEFALSPQGSFLEDTTGEQFLTYRYDDQTSRNTRS----------------------- 70
           | |:|    |  | :||    :     :|  |:    :|::
Db     471 FGPDFPHCPNTS-VEDTCVQLCARHRDSDPRICHTKNGSLIWADGTPCGPGHLCLDGSCV 529

Qy      71 -DED-       KDGNWDAWGDWSDCSRTCGGGASYSLRRCLT-     GRNCEGQNIRYK 116
             ||          ||:|  |   |   ||||||||  :| | |          || | |: ::|:
Db     530 LKEDVENPKAVVDGDWGPWRPWGQCSRTCGGGIQFSNRECDNPMPQNGGRFCLGERVKYQ 589

Qy     117 TCSNHDCPPDAEDFRAQQC---SAYNDVQYQGHYYEWLPRYN--DPAAPCALKCHAQGQN 171
           :|:   :|||: : || |||     :|||         |:: :|:|:|:     |     | | | | |:|::
Db     590 SCNTEECPPNGKSFREQQCEKYNAYNHTDLDGNFLQWVPKYSGVSPRDRCKLFCRARGRS 649

Qy     172 LVVELAPKVLDGTRCNTDSLDMCISGICQAVGCDRQLGSNAKEDNCGVCAGDGSTCRLVR 231
                   ||:||| |   |:| :|: | |    |||    : |   | | |||| | |: || :
Db     650 EFKVFEAKVIDGTLCGPETLSICVRGQCVKAGCDHVVNSPKKLDKCGVCGGKGTACRKIS 709

Qy     232 GQSKSHVSPEKREENVIAVPLGSRSVRITVKGPAHLFIESK----TLQGSKGEHSFNS-- 285
           |   |              ....:| |        | ||   |                |     |..  |
Db     710 G---SPTPFSYGYNDIVTIPAG--ATNIDVKQRSHPGVRNDGSYLALKTANGQYLLNGNL 764

Qy     286 -----PGVFVVENTTVEFQRGS----ER-QTFK-IPGPLMADFIF--------KTRYT-- 324
                     :|: | ::: ||      || |:|: :| ||        :              | |||
Db     765 AISAIEQDILVKGTILKYS-GSMATLERLQSFQALPEPLTVQLLTVSGEVFPPKVRYTFF 823

Qy     325 ---------AAKDSVVQFFFYQPISHQWRQTDFFPCTVTCGGGYQLNSAEC-------- 366
                         .|.                  | .|     |:   |   || |  .|    . ||
Db     824 VPNDMDFSVQNSKERATTNIIQSLPSAEWVIGDWSECPSTCRGSWQRRTVECRDPSGQAS 883

Qy     367 --VDIRLKRVVPDHYCHYYPENVKPKPKLKECSMDPCP 402
               |   ||               ||: ||        |    |||
Db     884 DTCDEALK            PEDAKF       CGSQFCF 904


RESULT 15
US-09-369-364A-2
; Sequence 2, Application US/09369364A
; Patent No. 6391610
; GENERAL INFORMATION:
;  APPLICANT:  Apte, Suneel
;  APPLICANT:  Hurskainen, Tiina L.
;  APPLICANT:  Hirohata, Satoshi
;  TITLE OF INVENTION:  Nucleic Acids Encoding Zinc Metalloproteases
;  FILE REFERENCE:  26473/4007/10-30-00
;  CURRENT APPLICATION NUMBER: US/09/369,364A
;  CURRENT FILING DATE:  1999-08-06
;  NUMBER OF SEQ ID NOS:  31
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 930
;   TYPE: PRT
;   ORGANISM: mus musculus ADAMTS-5
US-09-369-364A-2
```

```
Query Match             5.2%;  Score 482;  DB 4;  Length 930;
Best Local Similarity   29.5%;  Pred. No. 5.5e-28;
Matches  133;  Conservative   67;  Mismatches  170;  Indels  136;  Gaps   21;

Qy         34 FLPEEALSP                   QGSF          LEDI    IGEQFLIYRYDDQISRN 67
                 | ||:::  |                   |:            :| |      |     |   :  ::| .
Db        501 FGEEYSVCPGMDVCARIWCAVVRQGQMVCLIKKLPAVEGIPCGKCRVCLQCKCVDKIKKK 560

Qy         68 IRSDEDKIPCNWDAWCDWSPCSRIPCGGASYSLRRCP         LPGRNCEGQNIRYKIPSNH 121
                   |       |||  :||  ,   ..:||||   :: |  |           :||  ,  |:      ::|
Db        561 YYSCISSHGNWGSWGPWGQCSRSCGGGVQTAYRHCNNPAPRNGGRYCTGKRAIYRSCSVI 619

Qy        122 PCPPDAEDFRAQQCSAYNDVQYQGH        YYEWLPRYND   PAAPCALKCHAQGQNLVVE 175
                   |||:  :  ||  :||  |  |    |              : ||:|:|       ||    |  |  |  , ::|        |
Db        620 PCPPNGKSFRHEQCEAKNGYQSDAKGVKTFVEWVPKYAGVLPADVCKLTCRAKGTGYYVV 679

Qy        176 LAPKVLDGTRCNTDGLDMCISGICQAVGCDRQLGSNAKEDNCGVCAGDGSTCRLVRG     232
                  ::.|  |||  |      |    :   :   .   |       ||'    :·|      :  |  |:||  |    ::|    :  '
Db        680 FSPKVTDGTECRPYSNSVCVRGRCVRTGCDGIIGSKLQYDKCGVCGGDNSSCTKIIGTIN 739

Qy        233  QSKSHVSPEKREENVIAVPLGSRSVRI            IVKGPAHLFIESKTLQGSKGEHSLN 284
                  :||  :                 :|:  :.  |:   :::                :  ||:|  ::  ||         ||:    |
Db        740 KKSKGYT          DVVRIPEGATHIKVRQFKAKDQTRFTAYLALKKK      GEYLLN 787

Qy        285          SPGVFVVENTTVEFQRGSERQTFKIPGPLMADFIFKTRYTAAKD SVVQFFFY 336
                          |    :    :    |  .  :      |  |                   ||:       |:|  |:   :|·
Db        788 GKYMISTSETIIDINGTVMNYSGWSHRD             DFLHGMGYSATKEILIVQILAT 837

Qy        337 QPISHQWRQTDFFPCTVTPCGGGTQLNSADCVDIRLKRVVPDHYCHYYPENVKPKPKLKE  396
                 |                                      :  :  :|             |   :  |:        |
Db        838 DP                                  IKALGVR          YSPFVPKKISKVN       860

Qy        397 SMDPCPSSDGFKEIMPYDHFQPLPRWEHNPWTACSVSCGGGIQRRSFVCVE   ESMHGE 453
                          |  |     ::  :      |        |  :|      |: |||  :|     |         :  :     |   :           :      |
Db        861       SVISHGGNKVGI    HSTQL QWVTGPWLACSRTCDTGWHTRTVQCQDGNRKLAKG  912

Qy        454 ILQVEEWKCMYAPKPKVMQTCNLFDC 479
                                |:  :  :|        :  |          |
Db        913             CLLSQRPSAFKQCLLKKC 930

Search completed: March 22, 2003, 13:32:50
Job time : 38 secs
```

```
                    GenCore version 5.1.4 p5_4578
             Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on:          March 22, 2003, 13:30:49 ; Search time 31 Seconds
                                            (without alignments)
                                            2016.279 Million cell updates/sec

Title:           US-09-784-358-2
Perfect score:   9346
Sequence:        1 MASWISPWWVLIGMVFMHSP..........LNLCSLDYKQRCQSCQEI 1691

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        221153 seqs, 54462247 residues

Total number of hits satisfying chosen parameters:       221153

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries

Database :        Published Applications AA:*
                 1: [illegible] NEW_PUB.pep:*
                 2: [illegible] NEW_PUB.pep:*
                 3: [illegible] NEW_PUB.pep:*
                 4: [illegible] pep:*
                 5: [illegible] NEW_PUB.pep:*
                 6: [illegible] pep:*
                 7: [illegible] COMP.pep:*
                 8: [illegible] COMP.pep:*
                 9: [illegible] NEW_PUB.pep:*
                10: [illegible] pep:*
                11: [illegible] NEW_PUB.pep:*
                12: [illegible] pep:*
                13: [illegible] NEW_PUB.pep:*
                14: [illegible] pep:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                              SUMMARIES
              %
Result       Query
  No.  Score Match Length DB ID                             Description
------------------------------------------------------------------------------
    1   9346 100.0   1691 10 US-09-784-358-2                Sequence 2, Appli
    2   8882  95.0   1617 10 US-09-784-358-16               Sequence 16, Appl
    3   4689  50.2    845 10 US-09-784-358-12               Sequence 12, Appl
    4   4225  45.2    771 10 US-09-784-358-14               Sequence 14, Appl
    5 3956.5  42.3   1762  9 US-10-044-897-2               Sequence 2, Appli
    6   3956  42.3    724 10 US-09-784-358-8                Sequence 8, Appli
    7   3861  41.3   1745 10 US-09-800-729-89               Sequence 89, Appl
    8 3860.5  41.3   1766 10 US-09-764-853-554             Sequence 554, App
    9   3492  37.4    656 10 US-09-784-358-10               Sequence 10, Appl
   10   2891  30.9    874  9 US-09-796-753-20              Sequence 20, Appl
   11   2344  25.1    446 10 US-09-784-358-4                Sequence 4, Appli
   12   2058  22.0    525  9 US-09-978-295A-301            Sequence 301, App
   13   2058  22.0    525  9 US-09-978-697-301             Sequence 301, App
   14   2058  22.0    525  9 US-09-978-192A-301            Sequence 301, App
   15   2058  22.0    525  9 US-09-999-832A-301            Sequence 301, App
   16   2058  22.0    525  9 US-09-978-189-301             Sequence 301, App
   17   2058  22.0    525  9 US-10-174-590-436             Sequence 436, App
   18   2058  22.0    525  9 US-10-176-758-436             Sequence 436, App
   19   2058  22.0    525  9 US-10-175-737-436             Sequence 436, App
   20   2058  22.0    525  9 US-10-173-706-436             Sequence 436, App
   21   2058  22.0    525  9 US-10-175-738-436             Sequence 436, App
   22   2058  22.0    525  9 US-10-175-752-436             Sequence 436, App
   23   2058  22.0    525  9 US-10-176-482-436             Sequence 436, App
   24   2058  22.0    525  9 US-10-176-757-436             Sequence 436, App
   25   2058  22.0    525  9 US-10-176-913-436             Sequence 436, App
   26   2058  22.0    525  9 US-10-180-552-436             Sequence 436, App
   27   2058  22.0    525  9 US-10-180-557-436             Sequence 436, App
   28   2058  22.0    525  9 US-10-173-750-436             Sequence 436, App
   29   2058  22.0    525  9 US-10-174-572-436             Sequence 436, App
   30   2058  22.0    525  9 US-10-174-579-436             Sequence 436, App
   31   2058  22.0    525  9 US-10-174-582-436             Sequence 436, App
   32   2058  22.0    525  9 US-10-174-588-436             Sequence 436, App
   33   2058  22.0    525  9 US-10-175-739-436             Sequence 436, App
   34   2058  22.0    525  9 US-10-175-740-436             Sequence 436, App
   35   2058  22.0    525  9 US-10-175-743-436             Sequence 436, App
   36   2058  22.0    525  9 US-10-176-488-436             Sequence 436, App
   37   2058  22.0    525  9 US-10-176-492-436             Sequence 436, App
   38   2058  22.0    525  9 US-10-176-747-436             Sequence 436, App
   39   2058  22.0    525  9 US-10-176-750-436             Sequence 436, App
   40   2058  22.0    525  9 US-10-176-985-436             Sequence 436, App
   41   2058  22.0    525  9 US-10-176-987-436             Sequence 436, App
   42   2058  22.0    525  9 US-10-176-991-436             Sequence 436, App
   43   2058  22.0    525  9 US-10-176-992-436             Sequence 436, App
   44   2058  22.0    525  9 US-10-176-993-436             Sequence 436, App
   45   2058  22.0    525  9 US-10-184-658-436             Sequence 436, App

                              ALIGNMENTS

RESULT 1
US-09-784-358-2
; Sequence 2, Application US/09784358
; Patent No. US20020090627A1
; GENERAL INFORMATION:
;  APPLICANT: Donoho, Gregory
;  APPLICANT:  Scoville, John
;  APPLICANT:  Turner, C. Alexander Jr.
;  APPLICANT:  Friedrich, Glenn
;  APPLICANT:  Zambrowicz, Brian
;  APPLICANT:  Sands, Arthur T.
;  TITLE OF INVENTION:  NOVEL HUMAN [illegible] PROTEINS AND
;  TITLE OF INVENTION:  [illegible] SAME
;  FILE REFERENCE: LEX-0134-USA
;  CURRENT APPLICATION NUMBER: US/09/784,358
;  CURRENT FILING DATE:  2001-02-15
;  PRIOR APPLICATION NUMBER: US 60/183,282
;  PRIOR FILING DATE: 2000-02-17
;  NUMBER OF SEQ ID NOS: 17
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 1691
;   TYPE: PRT
;   ORGANISM: homo sapiens
US-09-784-358-2

  Query Match             100.0%;  Score 9346;  DB 10;  Length 1691;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1691;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MASWISPWWVLIGMVFMHSFLPQITAEKSPGAYFLLEEFALSPQGSFLEELLIGRQFLIYKY 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MASWISPWWVLIGMVFMHSFLPQITAEKSPGAYFLLEEFALSPQGSFLEELLIGRQFLIYKY 60

Qy         61 DDQISRNIRGDLEKISNWLAWLDWGDGRICGGAAYLGLRSLLDGRNCDGLNITYKICSN 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DDQISRNIRGDLEKISNWLAWLDWGDGRICGGAAYLGLRSLLDGRNCDGLNITYKICSN 120

Qy        121 HDCPPDAEDFRAQQCSAYNLVQYQGHYYEWLPRYNDFAAPCALKCHAQGQNLVVELAPKV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 HDCPPDAEDFRAQQCSAYNLVQYQGHYYEWLPRYNDFAAPCALKCHAQGQNLVVELAPKV 180
```

```
Qy        181 LDGTRCNTDSLDMCISGICQAVGCDRQLGSNAKEDNCGVCAGDGSTCRLVRGQSKSHVSP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 LDGTRCNTDSLDMCISGICQAVGCDRQLGSNAKEDNCGVCAGDGSTCRLVRGQSKSHVSP 240

Qy        241 EKPEENVIAVPLGSRSVRITVKGPAHLFIESKTLQGSKGEHSFNSPGVFVVENTTVEFQR 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 EKPEENVIAVPLGSRSVRITVKGPAHLFIESKTLQGSKGEHSFNSPGVFVVENTTVEFQR 300

Qy        301 GSERQTFKIPGPLMADFIFKTRYTAAKDSVVQFFFYQPISHQWRQTDFFPCTVTCGGGYQ 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GSERQTFKIPGPLMADFIFKTRYTAAKDSVVQFFFYQPISHQWRQTDFFPCTVTCGGGYQ 360

Qy        361 LNSAECVDIRLKRVVPDHYCHYYPENVKPKPKLKECSMDPCPSSDGFKEIMPYDHFQPLP 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LNSAECVDIRLKRVVPDHYCHYYPENVKPKPKLKECSMDPCPSSDGFKEIMPYDHFQPLP 420

Qy        421 RWEHNPWTACSVSCGGGIQRRSFVCVEESMHGEILQVEEWKCMYAPKPKVMQTCNLFDCP 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 RWEHNPWTACSVSCGGGIQRRSFVCVEESMHGEILQVEEWKCMYAPKPKVMQTCNLFDCP 480

Qy        481 KWIAMEWSQCTVTCGRGLRYRVVLCINHRGEHVGGCNPQLKLHIKEECVIPIPCYKPKEK 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 KWIAMEWSQCTVTCGRGLRYRVVLCINHRGEHVGGCNPQLKLHIKEECVIPIPCYKPKEK 540

Qy        541 SPVEAKLPWLKQAQELEETRIATEEPTFIPEPWSACSTTCGPCVQVREVKCRVLLTFTQT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 SPVEAKLPWLKQAQELEETRIATEEPTFIPEPWSACSTTCGPCVQVREVKCRVLLTFTQT 600

Qy        601 ETELPEEECEGPKLPTERPCLLEACDESPASRELDIPLPEDSETTYDWEYAGFTPCTATC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ETELPEEECEGPKLPTERPCLLEACDESPASRELDIPLPEDSETTYDWEYAGFTPCTATC 660

Qy        661 LGGHQEAIAVCLHIQTQQTVNDSLCDMVHRPPAMSQACNTEPCPPRWHVGSWGPCSATCG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 LGGHQEAIAVCLHIQTQQTVNDSLCDMVHRPPAMSQACNTEPCPPRWHVGSWGPCSATCG 720

Qy        721 VGIQTRDVYCLHPGETPAPPEECRDEKPHALQACNQFDCPPGWHIEEWQQCSRTCGGGTQ 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 VGIQTRDVYCLHPGETPAPPEECRDEKPHALQACNQFDCPPGWHIEEWQQCSRTCGGGTQ 780

Qy        781 NRRVTCRQLLTDGSFLNLSDELCQGPKASSHKSCARTDCPPHLAVGDWSKCSVSCGVGIQ 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 NRRVTCRQLLTDGSFLNLSDELCQGPKASSHKSCARTDCPPHLAVGDWSKCSVSCGVGIQ 840

Qy        841 RRKQVCQRLAAKGRRIPLSEMMCRDLPGFPLVRSCQMPECSKIKSEMKTKLGEQGPQILS 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 RRKQVCQRLAAKGRRIPLSEMMCRDLPGFPLVRSCQMPECSKIKSEMKTKLGEQGPQILS 900

Qy        901 VQRVYIQTRFEKRINLTIGSRAYLLPNTSVIIKCPVRRFQKSLIQWEKDGRCLQNSKRLG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 VQRVYIQTRFEKRINLTIGSRAYLLPNTSVIIKCPVRRFQKSLIQWEKDGRCLQNSKRLG 960

Qy        961 ITKSGSLKIHGLAAPDIGVYRCIAGSAQETVVLKLIGTDNRLIARPALREPMREYPGMDH 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 ITKSGSLKIHGLAAPDIGVYRCIAGSAQETVVLKLIGTDNRLIARPALREPMREYPGMDH 1020

Qy       1021 SEANSLGVTWHKMRQMWNNKNDLYLDDDHISNQPFLRALLGHCSNSAGSTNSWELKNKQF 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 SEANSLGVTWHKMRQMWNNKNDLYLDDDHISNQPFLRALLGHCSNSAGSTNSWELKNKQF 1080

Qy       1081 EAAVKQGAYSMDTAQFDELIRNMSQLMETGEVSDDLASQLIYQLVAELAKAQPTHMQWRG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 EAAVKQGAYSMDTAQFDELIRNMSQLMETGEVSDDLASQLIYQLVAELAKAQPTHMQWRG 1140

Qy       1141 IQEEIPPAAQLRGETGSVSQSSHAKNSGKLTFKPKGPVLMRQSQPPSISFNKTINSRIGN 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 IQEEIPPAAQLRGETGSVSQSSHAKNSGKLTFKPKGPVLMRQSQPPSISFNKTINSRIGN 1200

Qy       1201 TVYITKRTEVINILCDLITPSEATYTWTKDGTLLQPSVKIILDGTGKIQIQNPTRKEQGI 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TVYITKRTEVINILCDLITPSEATYTWTKDGTLLQPSVKIILDGTGKIQIQNPTRKEQGI 1260

Qy       1261 YECSVANHLGSDVESSSVLYAEAPVILSVERNITKPEHNHLSVVVGGIVEAALGANVTIR 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 YECSVANHLGSDVESSSVLYAEAPVILSVERNITKPEHNHLSVVVGGIVEAALGANVTIR 1320

Qy       1321 CPVKGVPQPNITWLKRGGSLSGNVSLLFNGSLLLQNVSLENEGTYVCIATNALGKAVATS 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CPVKGVPQPNITWLKRGGSLSGNVSLLFNGSLLLQNVSLENEGTYVCIATNALGKAVATS 1380

Qy       1381 VLHLLERRWPESRIVFLQGHKKYILQATNTRTNSNDPTGEPPPQEPFWEPGNWSHCSATC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 VLHLLERRWPESRIVFLQGHKKYILQATNTRTNSNDPTGEPPPQEPFWEPGNWSHCSATC 1440

Qy       1441 GHLGARIQRPQCVMANGQEVSEALCDHLQKPLAGFEPCNIRDCPARWFTSVWSQCSVSCG 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GHLGARIQRPQCVMANGQEVSEALCDHLQKPLAGFEPCNIRDCPARWFTSVWSQCSVSCG 1500

Qy       1501 EGYHSRQVTCKRTKANGTVQVVSPRACAPKDRPLGRKPCFGHPCVQWEPGNRCPGRCMGR 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 EGYHSRQVTCKRTKANGTVQVVSPRACAPKDRPLGRKPCFGHPCVQWEPGNRCPGRCMGR 1560

Qy       1561 AVRMQQRHTACQHNSSDSNCDDRKRPTLRRNCTSGACDVCWHTGPWKPCTAACGRGFQSR 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 AVRMQQRHTACQHNSSDSNCDDRKRPTLRRNCTSGACDVCWHTGPWKPCTAACGRGFQSR 1620

Qy       1621 KVDCIHTRSCKPVAKRHCVQKKKPISWRHCLGPSCDRDCTDTTHYCMFVKHLNLCSLDRY 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 KVDCIHTRSCKPVAKRHCVQKKKPISWRHCLGPSCDRDCTDTTHYCMFVKHLNLCSLDRY 1680

Qy       1681 KQRCCQSCQEG 1691
              |||||||||||
Db       1681 KQRCCQSCQEG 1691


RESULT 2
US-09-784-358-16
; Sequence 16, Application US/09784358
; Patent No. US20020099027A1
; GENERAL INFORMATION:
;  APPLICANT: Donoho, Gregory
;  APPLICANT:  Scoville, John
;  APPLICANT:  Turner, C. Alexander Jr.
;  APPLICANT:  Friedrich, Glenn
;  APPLICANT:  Zambrowicz, Brian
;  APPLICANT:  Sands, Arthur T.
;  TITLE OF INVENTION: NOVEL HUMAN THROMBOSPONDIN REPEAT PROTEINS AND
;  TITLE OF INVENTION:  POLYNUCLEOTIDES ENCODING THE SAME
;  FILE REFERENCE: LEX-0134-USA
;  CURRENT APPLICATION NUMBER: US/09/784,358
;  CURRENT FILING DATE:  2001-02-15
;  PRIOR APPLICATION NUMBER: US 60/183,282
;  PRIOR FILING DATE: 2000-02-17
;  NUMBER OF SEQ ID NOS: 17
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16
;   LENGTH: 1617
;   TYPE: PRT
;   ORGANISM: homo sapiens
US-09-784-358-16

  Query Match             95.6%;  Score 8892;  DB 16;  Length 1617;
  Best Local Similarity   95.6%;  Pred. No. 0;
  Matches 1617;  Conservative    0;  Mismatches    0;  Indels   74;  Gaps    1;

Qy          1 MASWTSPWWVLIGMVFMHSPLPQTTAEKSPGAYFLPEFALSPQGSFLEDTTGEQFLTYRY 60
              ||||||||||||||||||||||||||||||
Db          1 MASWTSPWWVLIGMVFMHSPLPQTTAEKSP------------------------------ 30

Qy         61 [illegible]RNCEEQNIIRYKFCGN 120
                                                          ||||||||||||||||
Db         31                                             RNCEEQNIIRYKFCGN 46

Qy        121 HDCPPDAPPTPAQQCSAYNDVQYQCHYYEWLPPYNDPAAPCALKCHAQGQNLVVELAPKV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         47 HDCPPDAPPTPAQQCSAYNDVQYQCHYYEWLPPYNDPAAPCALKCHAQGQNLVVELAPKV 106
```

```
Qy        181 LDGIRCNIDSLDMCISGICQAVFDDRQIGGNAKIDNCGVCAGDGSTCRLVRGQSKSHVSP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 LDGIRCNIDSLDMCISGICQAVFDDRQIGGNAKIDNCGVCAGDGSTCRLVRGQSKSHVSP 166

Qy        241 EKPEENVIAVPLGSRSVRIIVKGPAHLFIESKTLQGSKCEHSFNSPGVFVVENTIVEFQR 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 EKPEENVIAVPLGSRSVRIIVKGPAHLFIESKTLQGSKCEHSFNSPGVFVVENTIVEFQR 226

Qy        301 GSERQTFKIPGFLMADFIEKIRYIAAKDSVVQFFFYQPISHQWRQTDFFPCIVTCGGGYQ 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 GSERQTFKIPGFLMADFIEKIRYIAAKDSVVQFFFYQPISHQWRQTDFFPCIVTCGGGYQ 286

Qy        361 LNSAECVDIRLKRVVPDHYCHYYPENVKPKPKLKECSMDPCPSSDGFKEIMPYDHFQPLP 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 LNSAECVDIRLKRVVPDHYCHYYPENVKPKPKLKECSMDPCPSSDGFKEIMPYDHFQPLP 346

Qy        421 RWEHNPWTACSVSCGGGIQRRSFVCVEESMHGEIIQVEEWKCMYAPKPKVMQTCNLFDCP 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        347 RWEHNPWTACSVSCGGGIQRRSFVCVEESMHGEIIQVEEWKCMYAPKPKVMQTCNLFDCP 406

Qy        481 KWIAMEWSQCTVTCGRGLRYKVVLCINHRGEHVGGCNPQLKLHIKEECVIPIPCYKPKEK 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        407 KWIAMEWSQCTVTCGRGLRYKVVLCINHRGEHVGGCNPQLKLHIKEECVIPIPCYKPKEK 466

Qy        541 SPVEAKLPWLKQAQELEEIRIATEEPTFIPEPWSACSTTCGPGVQVREVKCRVLLIFTQT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        467 SPVEAKLPWLKQAQELEEIRIATEEPTFIPEPWSACSTTCGPGVQVREVKCRVLLIFTQT 526

Qy        601 EIELPEEECEGPKLPTERPCLLEACDESPASRELDIPLPEDSETTYDWEYAGFTPCTATC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        527 EIELPEEECEGPKLPTERPCLLEACDESPASRELDIPLPEDSETTYDWEYAGFTPCTATC 586

Qy        661 LGGHQEAIAVCLHIQTQQTVNDSLCDMVHRPPAMSQACNTEPCPPRWHVGSWGPCSATCG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        587 LGGHQEAIAVCLHIQTQQTVNDSLCDMVHRPPAMSQACNTEPCPPRWHVGSWGPCSATCG 646

Qy        721 VGIQTRDVYCLHPGETPAPPEECRDEKPHALQACNQFDCPPGWHIEEWQQCSRTCGGGTQ 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        647 VGIQTRDVYCLHPGETPAPPEECRDEKPHALQACNQFDCPPGWHIEEWQQCSRTCGGGTQ 706

Qy        781 NRRVTCRQLLTDGSFLNLSDELCQGPKASSHKSCARTDCPPHLAVGDWSKCSVSCGVGIQ 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        707 NRRVTCRQLLTDGSFLNLSDELCQGPKASSHKSCARTDCPPHLAVGDWSKCSVSCGVGIQ 766

Qy        841 RRKQVCQRLAAKGRRIPLSEMMCRDLPGFPLVRSCQMPECSKIKSEMKTKLGEQGPQILS 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        767 RRKQVCQRLAAKGRRIPLSEMMCRDLPGFPLVRSCQMPECSKIKSEMKTKLGEQGPQILS 826

Qy        901 VQRVYIQTREEKRINLTIGSRAYLLPNTSVIIKCPVRRFQKSLIQWEKDGRCLQNSKRLG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        827 VQRVYIQTREEKRINLTIGSRAYLLPNTSVIIKCPVRRFQKSLIQWEKDGRCLQNSKRLG 886

Qy        961 IIKSGSLKIHGLAAPDIGVYRCIAGSAQETVVLKLIGTDNRLIARPALREPMKEYPGMDH 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        887 IIKSGSLKIHGLAAPDIGVYRCIAGSAQETVVLKLIGTDNRLIARPALREPMKEYPGMDH 946

Qy       1021 SEANSLGVIWHKMPQMWNNKNLLYLGGHHLGNQPTLRALLGHTGNGAGTTNGWELKNKQF 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        947 SEANSLGVIWHKMPQMWNNKNLLYLGGHHLGNQPTLRALLGHTGNGAGTTNGWELKNKQF 1006

Qy       1081 EAAVKQGAYSMDTAQFDELIRNMSQLMETGEVSDDLASQLIYQLVAELAKAQPTHMQWRG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1007 EAAVKQGAYSMDTAQFDELIRNMSQLMETGEVSDDLASQLIYQLVAELAKAQPTHMQWRG 1066

Qy       1141 IQEETPPAAQLRGETGSVSQSSHAFNSTFLIIEKRKGEVLMLQQEPFIIENKTINGSRIN 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1067 IQEETPPAAQLRGETGSVSQSSHAFNSTFLIIEKRKGEVLMLQQEPFIIENKTINGSRIN 1126

Qy       1201 IVYITKRTEVINILCDLITPSEATYTWTKDGTLLQPSVKIISDESTFLQILNRIKREDSI 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1127 IVYITKRTEVINILCDLITPSEATYTWTKDGTLLQPSVKIISDESTFLQILNRIKREDSI 1186

Qy       1261 YECSVANHLGSDVESSSVIYAEAPVILSVERNIIKPEHNHLSVVVGGIVEAALGANVTIR 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1187 YECSVANHLGSDVESSSVIYAEAPVILSVERNIIKPEHNHLSVVVGGIVEAALGANVTIR 1246

Qy       1321 CPVKGVPQPNIIWLKRGGSLSGNVSLLFNGSLLLQNVSLENEGTYVCIATNALGKAVATS 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1247 CPVKGVPQPNIIWLKRGGSLSGNVSLLFNGSLLLQNVSLENEGTYVCIATNALGKAVATS 1306

Qy       1381 VLHLLERRWPESRIVFLQGHKKYILQAINTRTNSNDPTGEPPPQEPFWEPGNWSHCSATC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1307 VLHLLERRWPESRIVFLQGHKKYILQAINTRTNSNDPTGEPPPQEPFWEPGNWSHCSATC 1366

Qy       1441 GHLGARTQRPQCVMANGQEVSEALCDHLQKPLAGFEPCNIRDCPARWFTSVWSQCSVSCG 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1367 GHLGARTQRPQCVMANGQEVSEALCDHLQKPLAGFEPCNIRDCPARWFTSVWSQCSVSCG 1426

Qy       1501 EGYHSRQVTCKRTKANGTVQVVSPRACAPKDRPLGRKPCFGHPCVQWEPGNRCPGRCMGR 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1427 EGYHSRQVTCKRTKANGTVQVVSPRACAPKDRPLGRKPCFGHPCVQWEPGNRCPGRCMGR 1486

Qy       1561 AVRMQQRHTACQHNGSDSNCDDRKRPTLRRNCTSGACDVCWHTGPWKPCTAACGRGFQSR 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1487 AVRMQQRHTACQHNGSDSNCDDRKRPTLRRNCTSGACDVCWHTGPWKPCTAACGRGFQSR 1546

Qy       1621 KVECTHTRSCKPVAKRHCVQKKKPISWRHCLGPSCDRDCTDTTHYCMFVKHLNLCSLDRY 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1547 KVECTHTRSCKPVAKRHCVQKKKPISWRHCLGPSCDRDCTDTTHYCMFVKHLNLCSLDRY 1606

Qy       1681 KQRCCQSCQEG 1691
              |||||||||||
Db       1607 KQRCCQSCQEG 1617


RESULT 3
US-09-784-358-12
; Sequence 12, Application US/09784358
; Patent No. US20020099027A1
; GENERAL INFORMATION:
;  APPLICANT: Donoho, Gregory
;  APPLICANT:  Scoville, John
;  APPLICANT:  Turner, C. Alexander Jr.
;  APPLICANT:  Friedrich, Glenn
;  APPLICANT:  Zambrowicz, Brian
;  APPLICANT:  Sands, Arthur T.
;  TITLE OF INVENTION: NOVEL HUMAN THROMBOSPONDIN REPEAT PROTEINS AND
;  TITLE OF INVENTION:  POLYNUCLEOTIDES ENCODING THE SAME
;  FILE REFERENCE: LEX-0134 USA
;  CURRENT APPLICATION NUMBER: US/09/784,358
;  CURRENT FILING DATE:  2001-02-15
;  PRIOR APPLICATION NUMBER: US 60/183,282
;  PRIOR FILING DATE: 2000-02-17
;  NUMBER OF SEQ ID NOS: 17
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12
;   LENGTH: 845
;   TYPE: PRT
;   ORGANISM: homo sapiens
US-09-784-358-12

  Query Match             50.2%;  Score 4689;  DB 10;  Length 845;
  Best Local Similarity   100.0%;  Pred. No. 3.9e-259;
  Matches  830;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAGWISFWWVLLGMVFMESRLLQTIAKRSPGATFLLEEALSPQGSFLEDTIGEQFLIYRY 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAGWISFWWVLLGMVFMESRLLQTIAKRSPGATFLLEEALSPQGSFLEDTIGEQFLIYRY 60

Qy         61 DDQIGRNIRGGLDRKGNWGAWGDWSDCSRTCGGGASYSLRRCLTGRNCEGQNIRYKTCSN 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DDQIGRNIRGGLDRKGNWGAWGDWSDCSRTCGGGASYSLRRCLTGRNCEGQNIRYKTCSN 120

Qy        121 IDCPTDAEDFRAQQCGAYNDVQYQGHYYEWLPRYNDPAAPCALKCHAQGQNLVVELAPKV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 HDCPPDAEDFRAQQCSAYNDVQYQGHYYEWLPRYNDPAAPCALKCHAQGQNLVVELAPKV 180

Qy        181 LDGIRCNIDSLDMCISGICQAVGCDRQLGSNAKEDNCGVCAGDGSTCRLVRGQSKSHVSP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 LDGIRCNIDSLDMCISGICQAVGCDRQLGSNAKEDNCGVCAGDGSTCRLVRGQSKSHVSP 240

Qy        241 EKREENVIAVPLGSRSVRITVKGPAHLFIESKTLQGSKGEHSFNSPGVFVVENTTVEFQR 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 EKREENVIAVPLGSRSVRITVKGPAHLFIESKTLQGSKGEHSFNSPGVFVVENTTVEFQR 300

Qy        301 GSERQTFKIPGPLMADFIFKTRYTAAKDSVVQFFFYQPISHQWRQTDFFPCTVTCGGGYQ 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GSERQTFKIPGPLMADFIFKTRYTAAKDSVVQFFFYQPISHQWRQTDFFPCTVTCGGGYQ 360

Qy        361 LNSAECVDIRLKRVVPDHYCHYYPENVKPKPKLKECSMDPCPSSDGFKEIMPYDHFQPLP 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LNSAECVDIRLKRVVPDHYCHYYPENVKPKPKLKECSMDPCPSSDGFKEIMPYDHFQPLP 420

Qy        421 RWEHNPWTACSVSCGGGIQRRSFVCVEESMHGEIIQVEEWKCMYAPKPKVMQTCNLFDCP 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 RWEHNPWTACSVSCGGGIQRRSFVCVEESMHGEIIQVEEWKCMYAPKPKVMQTCNLFDCP 480

Qy        481 KWIAMEWSQCTVTCGRGLRYRVVLCINHRGEHVGGCNPQLKLHIKEECVIPIPCYKPKEK 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 KWIAMEWSQCTVTCGRGLRYRVVLCINHRGEHVGGCNPQLKLHIKEECVIPIPCYKPKEK 540

Qy        541 SPVEAKLPWLKQAQELEETRIATEEPTFIPEPWSACSTTCGPGVQVREVKCRVLLTFTQT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 SPVEAKLPWLKQAQELEETRIATEEPTFIPEPWSACSTTCGPGVQVREVKCRVLLTFTQT 600

Qy        601 ETELPEEECEGPKLPTERPCLLEACDESPASRELDIPLPEDSETTYDWEYAGFTPCTATC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ETELPEEECEGPKLPTERPCLLEACDESPASRELDIPLPEDSETTYDWEYAGFTPCTATC 660

Qy        661 LGGHQEAIAVCLHIQTQQTVNDSLCDMVHRPPAMSQACNTEPCPPRWHVGSWGPCSATCG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 LGGHQEAIAVCLHIQTQQTVNDSLCDMVHRPPAMSQACNTEPCPPRWHVGSWGPCSATCG 720

Qy        721 VGIQTRDVYCLHPGETPAPPEECRDEKPHALQACNQFDCPPGWHIEEWQQCSRTCGGGTQ 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 VGIQTRDVYCLHPGETPAPPEECRDEKPHALQACNQFDCPPGWHIEEWQQCSRTCGGGTQ 780

Qy        781 NRRVTCRQLLTDGSFLNLSDELCQGPKASSHKSCARTDCPPHLAVGDWSK 830
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 NRRVTCRQLLTDGSFLNLSDELCQGPKASSHKSCARTDCPPHLAVGDWSK 830
```

```
RESULT 4
US-09-784-358-14
; Sequence 14, Application US/09784358
; Patent No. US20020099027A1
; GENERAL INFORMATION:
;  APPLICANT: Donoho, Gregory
;  APPLICANT:  Scoville, John
;  APPLICANT:  Turner, C. Alexander Jr
;  APPLICANT:  Friedrich, Glenn
;  APPLICANT:  Zambrowicz, Brian
;  APPLICANT:  Sands, Arthur T.
;  TITLE OF INVENTION: NOVEL HUMAN THROMBOSPONDIN REPEAT PROTEINS AND
;  TITLE OF INVENTION.  POLYNUCLEOTIDES ENCODING THE SAME
;  FILE REFERENCE: LEX-0134-USA
;  CURRENT APPLICATION NUMBER: US/09/784,358
;  CURRENT FILING DATE:  2001-02-15
;  PRIOR APPLICATION NUMBER: US 60/183,282
;  PRIOR FILING DATE: 2000-02-17
;  NUMBER OF SEQ ID NOS: 17
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14
;   LENGTH: 771
;   TYPE: PRT
;   ORGANISM: homo sapiens
US-09-784-358-14
```

```
  Query Match             45.2%;  Score 4225;  DB 10;  Length 771;
  Best Local Similarity   91.1%;  Pred. No. 8.9e-233;
  Matches  756;  Conservative    0;  Mismatches    0;  Indels   74;  Gaps    1;

Qy          1 MASWTSPWWVLIGMVFMHSPLPQTTAEKSPGAYFLPEFALSPQGSFLEDTTGEQFLTYRY 60
              ||||||||||||||||||||||||||||||
Db          1 MASWTSPWWVLIGMVFMHSPLPQTTAEKSP------------------------------ 30

Qy         61 [illegible]GRNCDGQNIEYKTCSN 120
                                                          ||||||||||||||||
Db         31                                             GRNCDGQNIEYKTCSN 46

Qy        121 HDCPPDAEDFRAQQCSAYNDVQYQGHYYEWLPRYNDPAAPCALKCHAQGQNLVVELAPKV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         47 HDCPPDAEDFRAQQCSAYNDVQYQGHYYEWLPRYNDPAAPCALKCHAQGQNLVVELAPKV 106

Qy        181 LDGIRCNIDSLDMCISGICQAVGCDRQLGSNAKEDNCGVCAGDGSTCRLVRGQSKSHVSP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 LDGIRCNIDSLDMCISGICQAVGCDRQLGSNAKEDNCGVCAGDGSTCRLVRGQSKSHVSP 166

Qy        241 EKREENVIAVPLGSRSVRITVKGPAHLFIESKTLQGSKGEHSFNSPGVFVVENTTVEFQR 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 EKREENVIAVPLGSRSVRITVKGPAHLFIESKTLQGSKGEHSFNSPGVFVVENTTVEFQR 226

Qy        301 GSERQTFKIPGPLMADFIFKTRYTAAKDSVVQFFFYQPISHQWRQTDFFPCTVTCGGGYQ 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 GSERQTFKIPGPLMADFIFKTRYTAAKDSVVQFFFYQPISHQWRQTDFFPCIVTCGGGYQ 286

Qy        361 LNSAECVDIRLKRVVPDHYCHYYPENVKPKPKLKECSMDPCPSSDGFKEIMPYDHFQPLP 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 LNSAECVDIRLKRVVPDHYCHYYPENVKPKPKLKECSMDPCPSSDGFKEIMPYDHFQPLP 346

Qy        421 RWEHNPWTACSVSCGGGIQRRSFVCVEESMHGEIIQVEEWKCMYAPKPKVMQTCNLFDCP 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        347 RWEHNPWTACSVSCGGGIQRRSFVCVEESMHGEIIQVEEWKCMYAPKPKVMQTCNLFDCP 406

Qy        481 KWIAMEWSQCTVTCGRGLRYRVVLCINHRGEHVGGCNPQLKLHIKEECVIPIPCYKPKEK 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        407 KWIAMEWSQCTVTCGRGLRYRVVLCINHRGEHVGGCNPQLKLHIKEECVIPIPCYKPKEK 466

Qy        541 SPVEAKLPWLKQAQELEETRIATEEPTFIPEPWSACSTTCGPGVQVREVKCRVLLTFTQT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        467 SPVEAKLPWLKQAQELEETRIATEEPTFIPEPWSACSTTCGPGVQVREVKCRVLLTFTQT 526

Qy        601 ETELPEEECEGPKLPTERPCLLEACDESPASRELDIPLPEDSETTYDWEYAGFTPCTATC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        527 ETELPEEECEGPKLPTERPCLLEACDESPASRELDIPLPEDSETTYDWEYAGFTPCTATC 586

Qy        661 LGGHQEAIAVCLHIQTQQTVNDSLCDMVHRPPAMSQACNTEPCPPRWHVGSWGPCSATCG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        587 LGGHQEAIAVCLHIQTQQTVNDSLCDMVHRPPAMSQACNTEPCPPRWHVGSWGPCSATCG 646

Qy        721 VGIQTRDVYCLHPGETPAPPEECRDEKPHALQACNQFDCPPGWHIEEWQQCSRTCGGGTQ 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        647 VGIQTRDVYCLHPGETPAPPEECRDEKPHALQACNQFDCPPGWHIEEWQQCSRTCGGGTQ 706

Qy        781 NRRVTCRQLLTDGSFLNLSDELCQGPKASSHKSCARTDCPPHLAVGDWSK 830
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        707 NRRVTCRQLLTDGSFLNLSDELCQGPKASSHKSCARTDCPPHLAVGDWSK 756
```

```
RESULT 5
US-10-044-807-2
; Sequence 2, Application US/10044807
, Patent No. US20020165187A1
; GENERAL INFORMATION:
;  APPLICANT: Yu, Xuanchuan
;  APPLICANT:  Turner, C. Alexander Jr.
;  TITLE OF INVENTION: No. US20020165187Alel Human Protease and Polynucleotides Encod
;  FILE REFERENCE: LEX-0298-USA
;  CURRENT APPLICATION NUMBER. US/10/044,807
;  CURRENT FILING DATE:  2002-01-11
;  PRIOR APPLICATION NUMBER: US 60/261,684
;  PRIOR FILING DATE: 2001-01-11
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 1762
;   TYPE: PRT
;   ORGANISM: homo sapiens
US-10-044-807-2

  Query Match             42.3%;  Score 3956.5;  DB 9;  Length 1762;
  Best Local Similarity   43.3%;  Pred. No. 4.7e-217;
  Matches  769;  Conservative  265;  Mismatches  554;  Indels  187;  Gaps   31;

Qy         64 ISRNTRSDEDKIGNWDAWGDWGDCGRTCGGGASYSLRRCLGGRNCEGQNIRYKTCSNHDC 123
               :||   ||:||:|| ||||| ||:||||||||||| ||||||: ::|||:|||:||||  :|
Db         22 SSRTARSEEDRIGLWDAWGPWSECSRTCGGGASYSLRRCLSGKSCEGRNIRYKTCSNVDC 81

Qy        124 PPDAEDFRAQQCSAYNDVQYQGHYYEWLPRYNDPAAPCALKCHAQGQNLVVELAPKVLDG 183
              ||:| ||||||||||:|||: | :||||  ||    ||:||| :|   |||| ||||||
Db         82 PPEAGDFRAQQCSAHNDVKHHGQFYEWLPVSNDPDNPCSLKCQAKGTTLVVELAPKVLDG 141

Qy        184 TRCNTDSLDMCISGICQAVGCDRQLGSNAKEDNCGVCAGDGSTCRLVRGQSKSHVSPEKR 243
              |||  |:|||||||||:||  |||  ||||    |||| |||  ||||||||||||| ||  :|   |
Db        142 TRCYTESLDMCISGLCQIVGCDHQLGSTVKEDNCGVCNGDGSTCRLVRGQYKSQLSAIKS 201

Qy        244 EENVIAVPLGSRSVRITVKGPAHLFIESKTLQGSKGEHSFNSPGVFVVENTTVEFQRGSE 303
              :: |:|:| |||  :|:  :|||  ||::|:||||||:||| :|  :| | |::::|:|:|:   :
Db        202 DDTVVAIPYGSRHIRLVLKGPDHLYLETKTLQGTKGENSLSSTGTFLVDNSSVDFQKFPD 261

Qy        304 RQTFKIPGPLMADFIFKTRYTAAKDSVVQFFFYQPISHQWRQTDFFPCTVTCGGGYQLNS 363
              ::   :: ||| |:|| | |  :  : || :|| || || |:||:|||||||: :||||:||  |
Db        262 KEILRMAGPLTADFIVKIRNSGSADSTVQFIFYQPIIHRWRETDFFPCSATCGGGYQLTS 321

Qy        364 AECVDIRLKRVVPDHYCHYYPENVKPKPKLKECSMDPCPSSDGFKEIMPYDHFQPLPRWE 423
              ||| |:|  ||| | ||||||||:|||||:|:|| :|||||:|:|:::||||| : |:| | |
Db        322 AECYDLRSNRVVADQYCHYYPENIKPKPKLQECNLDPCPASDGYKQIMPYDLYHPLPRWE 381

Qy        424 HNPWTACSVSCGGGIQRRSFVCVEESMHGEILQVEEWKCMYAPKPKVMQTCNLFDCPKWI 483
                 |||||| |||||||  |:    |||  :  | :   ||||||||| ||   : |  |::||||::
Db        382 ATPWTACSSSCGGGIQSRAVSCVEEDIQGHVTSVEEWKCMYTPKMPIAQPCNIFDCPKWL 441

Qy        484 AMEWSQCTVTCGRGLRYRVVLCINHRGEHVGGCNPQLKLHIKEECVIPIPCYKPKEKSPV 543
              | |||  |||||:||||||||||:||| |  |||:|: | |||||||::| ||||||| ||   |
Db        442 AQEWSPCTVTCGQGLRYRVVLCIDHRGMHTGGCSPKTKPHIKEECIVPTPCYKPKEKLPV 501

Qy        544 EAKLPWLKQAQELEETRIATEEPTFIPEPWSACSTTCGPGVQVREVKCRVLLTFTQTETE 603
              |||||| ||||||||       :|||:|||| ||||: ||| | ||: |:|:| |:::::  :
Db        502 EAKLPWFKQAQELEEGAAVSEEPSFIPEAWSACTVTCGVGTQVRIVRCQVLLSFSQSVAD 561

Qy        604 LPEEECEGPKLPTERPCLLEACDESPASRELDIPLP                EDSETTYDWEYAGF 653
              || :||||||  ::|          :|   | | |:         |         :|  :   ||||  |
Db        562 LPIDECEGPKPASQR          ACYAGPCSGEIPEFNPDETDGLFGGLQDFDELYDWEYEGF 616

Qy        654 TPCTATCGGGHQEATAVCLHIQTQQTVNDSLCDMVHRPPAMSQACNTEPCPPRWHVGSWG 713
              : :: :| || |:|: ||:  |::    ::||      ||| : ::|| :| | : | :|
Db        617 TKCSESCGGGVQEAVVSCLNKQTREPAEENLCVTSRRPPQLLKSCNLDPCPARWEIGKWS 676

Qy        714 PCSATCGVGIQTRDVYCLH         PGETPAPPEECRDEKPHALQACNQFDCPPGWHIEE 767
              ||| ||||||:|||||:|          ||      :| ||  ||  :|||:|:||||| |:   :
Db        677 PCSLTCGVGLQTRDVFCSHLLSREMNETVILADELCRQPKPSTVQACNRFNCPPAWYPAQ 736

Qy        768 WQQCSRTCGGGTQNRRVTCRQLLTDGSFLNLSDELCQGPKASSHKSCARTDCPPHLAVGD 827
              || |||||||| | | | |:| : |||||| | :  |    | :  ::| : || :     :|
Db        737 WQPCSRTCGGGVQKREVLCKQRMADGSFLELPETFCSASKPACQQACKKDDCPSEWLLSD 796

Qy        828 WSKCSVSCGVGIQRRKQVCQRLAAKGRRIPLSEMMCRDLPGFPLVRSCQMPECSKIKSEM 887
              | :|| |||:|| |  | :|:::  |      ::  :|  ||     :| | :   |::
Db        797 WTECSTSCGEGTQTRSAICRKMLKTGLSTVVNSTLCPPLPFSSSIRPCMLATCARPGRPS 856

Qy        888 KTKLGEQGPQILSVQRVYIQTRDEKRINLTIGSRAYLLPNTSVIIKCPVRRFQKSLIQWE 947
               ||       | | : :|||:|: :::::  :| |||||| |:|:::|| || : :| |  |
Db        856 STK      HSPHIAAARKVYIQTRRQRKLHFVVGGFAYLLPKTAVVLRCPARRVRKPLITWE 912

Qy        948 KDGRCLQNSKRLGITKSGSLKIHGLAAPDIGVYRCIAGSAQETVVLKLIGTDNRLIARP 1006
              :||: | :   : :      | | :| |  | ||  | ||: :    :|||| : :|:||
Db        913 KDGQHLISSTHVTVAPFGYLKIHRLKPSDAGVYTCSAGPAREHFVIKLIGGNRKLVARPL 972

Qy       1007                ALREPMKEYPGMSHGEANGLGVTWHK   MRQMWNNKNDLYLDDDHISNQPF 1055
                             | |:  :      |   | :         |    | : |     |   || :|
Db        973 SPRSEEEVLAGRKGGPKEALQTHKHQNGIFSNGSKAEKRGLAANPGSRY   DDLVSR 1027

Qy       1056 LRALLGHCSNSAGSINSWELKNKQFEAAVKQ       GA          YSMDTAQ   FDELIRN 1102
                   ||            |||  ::           :   ||          :::  |  |   |:::  |
Db       1028      LLEQGGWPGELLASWEAQDSAERNTTSEEDPGAEQVLLHLPFTMVTEQRRLDDILGN 1084

Qy       1103 MSQLMETGEVSDDLASQLIYQLVAELAKAQPTHMQ   WRGIQEETPP                   A 1148
              :||   |   |: |   :   |: ||   |: ::   |            :   :   ||                   :
Db       1085 LSQQPE   ELRDLYSKHLVAQLAQEIFRSHLEHQDTLLKPSERRTSPVTLSPHKHVSGFS 1142

Qy       1149 AQLR       GETGSVSQSSHAKNSGKLTFKPKGPVLMRQSQPPSISFNKTINSRIGNTVY 1203
              : ||        |: |  |:  | |                     | ::|:   :  :|  :: :  :  : :| ||
Db       1143 SSLRTSSTGDAGGGSRRPHRK                     PTILRKISAAQQLSASEVV  THLGQTVA 1190

Qy       1204 ITKRTEVINILCDLITPSEATYTWTKDGTLLQPSVKIIPDGPKIQIQNPTRKEQGIYEC 1263
              :   |  : : |: |       | :| ::|  :| | |:|:        :||  |    : | | |
Db       1191 LASGTLSVLLHCEAIGHPRPTISWARNGEEVQFSDRILLQPDDSLQILAPVEADVGFYTC 1250

Qy       1264 SVANHLGSDVESSSVLYALAPVTLSVEENITKPLHNHLSVVVGSIVLAAIGANVTIRCPV 1323
              :  | || |  | :| |  |::: :     :    |   ::  :|  ::    | |||| | |
Db       1251 NATNALGYDSVSIAVTLAGKPLVKTSRMTVINTEKPAVTVDIGSTIKTVQGVNVTINCQV 1310

Qy       1324 KGVPQPNTTWLKRGGSLSGNVSLLFNGSLLLQNVSLENEGTYVCIATNALGKAVATSVLH 1383
               |||:  :||  :     | |:    |   ||||| |||   ::| | | | |   |:   :: |
Db       1311 AGVPEAEVTWFRNKSKL GSPHHLHEGSLLLTNVSSSDQGLYSCRAANLHGELTESTQLL 1369

Qy       1384 LLERRWPESRIVPTQGHKKYILQATNTRTNS                  N P             TGEPP 1422
              :|:          ::: ::      :| ||       |                  ||             | |
Db       1370 ILDPPQVPTQLEDIRA       LLAATGPNLPSVLTSPLGTQLVLDPGNSALLGCPIKGHPV 1425

Qy       1423 P----      ----- --          -                                  1423
              |
Db       1426 PNITWFHGGQPIVTATGLTHHILAAGQILQVANLSGGSQGEFSCLAQNEAGVLMQKASLV 1485

Qy       1424 TQEPPWEPGNWSHCSATCGHLGARIQRPQCVMANGQEVSEALCDHLQKPLAG FEPCNIRD 1482
                |: :|          : |||:||:  | :    | :|::  |   ||: |  |        :|             ||  ||
Db       1486 TQDYWWSVDRLATCSASCGNRGVQQPRLRCLLNSTEVNPAHCAGKVRPAVQPIACNRRD 1544

Qy       1483 CPARWFTSVWSQCSVSCGEGYHSRQVTCKRTKANGTVQVVSPRACAP KDRPLGRKPCEG 1541
              ||:||   : || |: ||| |   :|:|||:: ||:|       ||     |          ||:   :  |
Db       1545 CPSRWMVTSWSACTRSCGGGVQTRRVTCQKLKASGISTPVSNDMCTQVAKRPVDTQACNQ 1604

Qy       1542 HPCVQW--EPGNRCPGRCMGRAVRMQQRHTACQHNSS       DSNCDPRKRPTLRRNCTSG 1595
                 ||:|          :| | ::|   : :| |    ||                   |     ||    :|| |
Db       1605 QLCVEWAFSSWGQCNGPCIGPHLAVQHRQVFCQTRDGITLPSEQCSALPRPVSTQNCWSE 1664

Qy       1596 ACDVCWHTGPWKPCTAACGR GFQSRKVDCIHTRSCKPVAKRHCVQKKKPISWRHCLGP 1653
              || | |   |   |||  ||  ||||:|:|:| |: | |  :   |       :|  :|:  |  :  |
Db       1665 ACSVHWRVSLWTLCTATCGNYGFQSRRVECVHARTNKAVPEHLCSWGPRPANWQRCNITP 1724

Qy       1654 SCDRDCTDTTHYCMFVKHLNLCSLDRYKQRCCQSC 1688
                : :| ||| ||  || | ||  : ::| ||  :|
Db       1725 CENMECRDTTRYCEKVKQLKLCQLSQFKSRCCGTC 1759


RESULT 6
US-09-784-358-8
; Sequence 8, Application US/09784358
; Patent No. US20020099027A1
; GENERAL INFORMATION:
;  APPLICANT: Donoho, Gregory
;  APPLICANT:  Scoville, John
;  APPLICANT:  Turner, C. Alexander Jr.
;  APPLICANT:  Friedrich, Glenn
;  APPLICANT:  Zambrowicz, Brian
;  APPLICANT:  Sands, Arthur T.
;  TITLE OF INVENTION: NOVEL HUMAN THROMBOSPONDIN REPEAT PROTEINS AND
;  TITLE OF INVENTION:  POLYNUCLEOTIDES ENCODING THE SAME
;  FILE REFERENCE: LEX-0134-USA
;  CURRENT APPLICATION NUMBER: US/09/784,358
;  CURRENT FILING DATE:  2001-02-15
;  PRIOR APPLICATION NUMBER: US 60/183,282
;  PRIOR FILING DATE: 2000-02-17
;  NUMBER OF SEQ ID NOS: 17
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8
;   LENGTH: 724
;   TYPE: PRT
;   ORGANISM: homo sapiens
US-09-784-358-8

  Query Match             42.3%;  Score 3956;  DB 10;  Length 724;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-217;
  Matches  706;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MASWISPWWVLIGMVFMHSPLLQTTAEKSPGAYFLPEFALSPQGSFLEDTTGEQFLTYRY 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MASWISPWWVLIGMVFMHSILLQTTAEKSPGAYFLPEFALSPQGSFLEDTTGEQFLTYRY 60

Qy         61 DDQTSENTPSDHDKPCNWLAWGDWSDCSETCGGGASYSLRRCLTGRNCEGQNIRYKTCSN 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DDQTSENTPSDHDKPCNWLAWGDWSDCSETCGGGASYSLRRCLTGRNCEGQNIRYKTCSN 120

Qy        121 HDCPPDAEDFRAQQCSAYNDVQYQGHYYEWLPRYNDPAAPCALKCHAQGQNLVVELAPKV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 HDCPPDAEDFRAQQCSAYNDVQYQGHYYEWLPRYNDPAAPCALKCHAQGQNLVVELAPKV 180

Qy        181 LDGTRCNTDSLDMCISGICQAVGCDRQLGSNAKEDNCGVCAGDGSTCRLVRGQSKSHVSP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 LDGTRCNTDSLDMCISGICQAVGCDRQLGSNAKEDNCGVCAGDGSTCRLVRGQSKSHVSP 240

Qy        241 EKRFENVIAVPLGSRSVRITVKGPAHLFIESKTLQGSKGEHSFNSPGVFVVENTTVEFQR 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 EKRFENVIAVPLGSRSVRITVKGPAHLFIESKTLQGSKGEHSFNSPGVFVVENTTVEFQR 300

Qy        301 GSERQTFKIPGPLMADFIFKTRYTAAKDSVVQFFFYQPISHQWRQTDFFPCTVTCGGGYQ 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GSERQTFKIPGPLMADFIFKTRYTAAKDSVVQFFFYQPISHQWRQTDFFPCTVTCGGGYQ 360

Qy        361 LNSAECVDIRLKRVVPDHYCHYYPENVKPKPKLKECSMDPCPSSDGFKEIMPYDHFQPLP 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LNSAECVDIRLKRVVPDHYCHYYPENVKPKPKLKECSMDPCPSSDGFKEIMPYDHFQPLP 420

Qy        421 RWEHNPWTACSVSCGGGIQRRSFVCVEESMHGEILQVEEWKCMYAPKPKVMQTCNLFDCP 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 RWEHNPWTACSVSCGGGIQRRSFVCVEESMHGEILQVEEWKCMYAPKPKVMQTCNLFDCP 480

Qy        481 KWIAMEWSQCTVTCGRGLRYRVVLCINHRGEHVGGCNPQLKLHIKEECVIPIPCYKPKEK 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 KWIAMEWSQCTVTCGRGLRYRVVLCINHRGEHVGGCNPQLKLHIKEECVIPIPCYKPKEK 540

Qy        541 SPVEAKLPWLKQAQELEETRIATEEPTFIPEPWSACSTTCGPGVQVREVKCRVLLTFTQT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 SPVEAKLPWLKQAQELEETRIATEEPTFIPEPWSACSTTCGPGVQVREVKCRVLLTFTQT 600

Qy        601 ETELPEEECEGPKLPTERPCLLEACDESPASRELDIPLPEDSETTYDWEYAGFTPCTATC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ETELPEEECEGPKLPTERPCLLEACDESPASRELDIPLPEDSETTYDWEYAGFTPCTATC 660

Qy        661 LGGHQEAIAVCLHIQTQQIVNDSLCDMVHRPPAMSQACNTEPCPPR 706
              ||||||||||||||||||||||||||||||||||||||||||||||
Db        661 LGGHQEAIAVCLHIQTQQIVNDSLCDMVHRPPAMSQACNTEPCPPR 706


RESULT 7
US-09-800-729-89
; Sequence 89, Application US/09800729
; Patent No. US20020068319A1
; GENERAL INFORMATION:
;  APPLICANT:  Ni et al.
;  TITLE OF INVENTION:  32 Human secreted proteins
;  FILE REFERENCE:  PZ044P1
;  CURRENT APPLICATION NUMBER: US/09/800,729
;  CURRENT FILING DATE:  2001-03-08
;  PRIOR APPLICATION NUMBER:  PCT/US00/26013
;  PRIOR FILING DATE:  2000-09-22
;  PRIOR APPLICATION NUMBER:  60/155,709
;  PRIOR FILING DATE:  1999-09-24
;  NUMBER OF SEQ ID NOS:  217
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 89
;   LENGTH: 1745
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-800-729-89

  Query Match             41.3%;  Score 3861;  DB 10;  Length 1745;
  Best Local Similarity   42.6%;  Pred. No. 1.3e-211;
  Matches  750;  Conservative  205;  Mismatches  552;  Indels  241;  Gaps   32;

Qy         64 TSENTPSDHDKPCNWLAWGDWSDCSETCGGGASYSLRRCLTGRNCEGQNIRYKTCSNHDC 123
              :||   ||:||:|| ||||| ||:|||||||||||||||||||: ::|||:||||:|||| ||
Db         22 GSKIAKEELDKLGLWLAWGEWSLCSRTCGGGATYSLRRCLSQKSCEGRNIKYKTCGNVDC 81

Qy        124 PPDAEDFRAQQCSAYNDVQYQGHYYEWLPRYNDPAAPCALKCHAQGQNLVVELAPKVLDG 183
              ||:| ||||||||||:|||:: | :||||||   |||   ||:||| |:|   ||||||||||||
Db         82 PFEAGSFRAQQCSAHNDVKHHGQFYEWLPVSNDPDNPCSLKCQAKGTTLVVELAPKVLDG 141

Qy        184 TRCNTDSLDMCISGICQAVGCDRQLGSNAKEDNCGVCAGDGSTCRLVRGQSKSHVSPEKR 243
              |||  |:|||||||||||:|| |||| ||||   |||||||| |||||||||||| || :| |
Db        142 TRCYTESLDMCISGECQIVGCDHQLGSTVKEDNCGVCNGDGSTCRLVRGQYKSQLSATKS 201

Qy        244 FENVIAVPLGSRSVRITVKGPAHLFIESKTLQGSKGEHSFNSPGVFVVENTTVEFQRGSE 303
              :: |:|:| ||| :|: :||| ||::|:|||||:|||:| :| | |:|:|::|:|||:  :
Db        202 DDTVVAIPYGSRHIRLVLKGPDHLYLETKTLQGTKGENSLSSTGTFLVDNSSVDFQKFPD 261

Qy        304 RQTFKIPGPLMADFIFKTRYTAAKDSVVQFFFYQPISHQWRQTDFFPCTVTCGGGYQLNS 363
              ::  :: ||| |||| | |  :  : || ||| ||||| |:||:||||||: |||||||| |
Db        262 KEILRMAGPLTADFIVKIRNSGSALSTVQYIFYQPIIHRWRETDFFPCSATCGGGYQLTS 321

Qy        364 AECVDIRLKRVVPDHYCHYYPENVKPKPKLKECSMDPCPSSDGFKEIMPYDHFQPLPRWE 423
              ||| |:|  ||| | |||||||||:|||||:||:|::||||:                 |||
Db        322 AECYDLRSNRVVADQYCHYYPENIKPKPKLQECNLDPCPA-----------------KWL 364

Qy        424 HNPWTACSVSCGGGIQRRSFVCVEESMHGEILQVEEWKCMYAPKPKVMQTCNLFDCPKWI 483
                |||||| |||||||| |:   |||| : | :  |||||||| ||  : | ||:||||||:
Db        365 APPWTACSSSCGGGIQSRAVSCVEEDIQGHVTSVEEWKCMYTPKMPIAQPCNIFDCPKWL 424

Qy        484 AMEWSQCTVTCGRGLRYRVVLCINHRGEHVGGCNPQLKLHIKEECVIPIPCYKPKEKSPV 543
              | ||| ||||||:||||||||||:||| | |||:|: | ||||||::| |||||||| ||
Db        425 AQEWSPCSVTCGQGLRYRVVLCIDHRGMHTGGCSPKTKPHIKEECIVPTPCYKPKEKLPV 484

Qy        544 EAKLPWLKQAQELEETRIATEEPTFIPEPWSACSTTCGPGVQVREVKCRVLLTFTQTETE 603
              |||||| ||||||||      :|||:|||: ||||: ||| | ||| |:|:|||:|:|:  :
Db        485 EAKLPWFKQAQELEEGAAVSEEPSFIPEAWSACTVTCGVGTQVRIVRCQVLLSFSQSVAD 544

Qy        604 LPEEECEGPKLPTERPCLLEACDESPASRELDIPLP----------EDSETTYDWEYAGF 653
              || :||||||   ::|         ||   | | |:    |             :| :  ||||| ||
Db        545 LPIDECEGPKPASQR     ACYAGPCSGEIPEFNPDETDGLFGGLQDFDELYDWEYEGF 599

Qy        654 TPCTATCLGGHQEAIAVCLHIQTQQIVNDSLCDMVHRPPAMSQACNTEPCPPRWHVGSWG 713
              | |: :| || |||:  ||: ||::     ::|(      ||| :  ::|| :||| || :| |
Db        600 TECSESCGGGVQEAVVYCLNKQIRLEPAEENLCVESRPLQLLKPCNLQPCPAEWLFGRWS 659

Qy        714 PCSATCGVGIQTRDVYCLH-----PCETPAPPEE-CRDEKPHALQACNQFDCPPGWHIEE 767
              ||| |||||:|||||:| |              |·       :| ||   ||   :|||||:|:||| |:   :
Db        660 PCSLTCGVGLQTRDVFCSHLLSREMNETVILADELCRQPKPSTVQACNRFNCPPAWYPAQ 719

Qy        768 WQQTSPTCGGGTQNRRVTCRQLLTDGSFLNLSDEEDQGPEASSHESCARTDCPPHLAVGD 827
              || ||||||||| | | | |:| : ||||| | :  |   | :  ::| : |||     : |
Db        720 WQPCSRTCGGGVQRRKVTCRQRMADGSFLELPETFCSASKPACQQACKKDDCPSEWLLSD 779

Qy        828 WSKCSVSCGVGIQRRKQVCQRLAAKGRRIPLSEMMCRDLPGFPLVRSCQMPECSKIKSEM 887
```

```
                 | :|| ||| | | |  :;:::    |     ::  :|   ||     :| | :  :::
Db        786 WIEQSESQGEQIQTESALRKKMLKIQLSIVVNQIDTPELEPSSSIKPQMLATAKTPQEE 838

Qy        888 KIKLGEQGPQTESVQRVVTQIEPEPPPINCLESSPAVLTENITVTTETVKEEQEKTLQWE 947
                 ||      | | : ::||||,| :::::  :|   ||||| |:|:::|| || :| || :|
Db        839 SIK   HSPHIAAARKVYTQIRRQRKLHEVVQGEAYLLPKIAVVLPQPAPRVEKPLTIWE 895

Qy        948 KDGRGLQNSKPLSLIKSTNLKIHGLAAPTGVYECIACSAQEIVVLKLIGTDNRLIAEP 1006
              |||: | :|   : :    | |,|| |    | ||, | || |:|  |:|||| : :|:||,
Db        896 KDGQHLTSTHVTVAPEGYLKIHHLKPSPAGVYTCSAGPAREHEVTKLTGNRKLVARPL 955

Qy       1007            ALREEPMEEYPGMDHSEANSLGVIWHK  MRQMWNNKNDLYLDDDHISNQPP 1055
                         | |:   :      |    | :      |    | :  |     | || :|
Db        956 SPKSEEEVLAGRKGGPKEAIQTHKHQNGTESNGSKAPKRGLAANPGSRY  PDLVSR    1010

Qy       1056 LRALGHTGNSAGSINSWELKNEQPEAAVKQ    GA          YGMPIAQ  PEPLTEN 1112
                 :|           ,||  :           :   ||          ::|  :  '   |:::,
Db       1011    LEEQSGWPGELLASWEAQDSAERNTTSEEDPGAEQVLLHLPFIMVTEQRRLDDTLGN 1067

Qy       1113 MSQLMEDGEVSDDLASQLIYQIVAELAKAQPTHMQ  WRGIQEEIPP          ---A 1148
               :||   |   |:  |   :   |: |,  |: ::     |           :   :  | '       :
Db       1068 LSQQPE  ELRDLYSKHLVAQLAQETFRSHLEHQDTLLKPSERRISPVILSPHKHVSGES 1125

Qy       1149 AQLR     GETGSVSQSSHAKNSGKLIFKPKGPVLMRQ SQPPSIGPNKIINGRIGNIVY 1203
               : ||      |: |   |: |   | :                  | ::|: |    :|  :  :  : :|  |
Db       1126 SSLRTSSIGDAGGGSRRPHRK                  PTTLRKISAAQQLSASEVV-IHLGQIVA 1173

Qy       1204 IIKRTEVINITGPITESEATYIWIKDGILLQPSVKIILDGTGKIQIQNPIRKEQGIYEC 1263
               :    :  :  :  |:|  |      ' :' ::|    :|  |  |  :|:|        :|'   |      :  |  |  |
Db       1174 LASGIISVLLHTEATSHPPPTISWARNGEEVQPSDRILQPDDSLQILAPVEADVGFYTC 1233

Qy       1264 SVANEKGSDVESSGVLYAEAPVILSVERNITKPEHNHLSVVVGGIVEAALGANVTIRCPV 1323
               .  | ||  |   |  :|   |   |::  :      :      |       ::|  :|    ::      | |:|| |  |
Db       1234 NAINAGSYTGVSTAVTLAGKPLVKTSRMTVINTEKPAVTVDTGSTIKIVQGVNVTINCQV 1293

Qy       1324 KGVPQPNITWLKRGGSLSGNVSLLFNGSLLLQNVSLENEGTYVCIATNALGKAVATSVLH 1383
                |||.   :||  :      |  |:     |     ||||| |||     ::|  |   |   |   |    |:      ::  |
Db       1294 AGVPEAEVTWFRNKSKLGSSPHHLHEGSLLLTNVSSSDQGLYSCRAANLHGELTESTQLL 1352

Qy       1384 ILEPPWPESRIVPLQGHKKYILQAINIRINS                NDP              IGEEP 1422
               '|:         :::  :  .         :|  ||            |                  ||             | |
Db       1353 ILDPPQVPTQLEDIRA       LLAATGPNLPSVLLSPLGTQLVLDPGNSALLGCPIKGHPV 1408

Qy       1423 P                                                          1423
              |
Db       1409 PNIIWFHGGQPIVTATGLTHHILAAGQILQVANLSGGSQGEFSCLAQNEAGVLMQKASLV 1468

Qy       1424  QEPPWPTGNWSHGSALGSHLSAETQRPGTVMANGQEVSEALGTHGQKELAGEEPCNIEG 1482
                ,  :       :  |||::|:  ,  :    |  :,::   |    '|:  | |          :.            '
Db       1469 IQDYWWSVDRLATGSASGGNKTVQQPRLRGLL NSTEVNPAHCAGKVREAVQPTACNKED 1527

Qy       1483 CPARWETSVWSQGSVSGGEGYHSGQVTGKKTKANGTVQVVSPEAGAP KDRPLGRKPGPT 1541
               || ,|   ' ||  |: |:|       :| ;|||:  ||:|           ||        |               : :  :
Db       1528 CPSRWMVTSWSACTESGGGGVQTRRVTCQKLKASGISTPVSNDMCTQVAKRPVDTQACNQ 1587

Qy       1542 HPCVQW  EPGNRGPGRCMGRAVRMQQRHTACQHNSS        DSNCDDRKRPTLRRNCTST 1595
                 ||:|            :!  |   |:'    :  :|   |        ;|                       |         ||       :||  |
Db       1588 QLCVEWAFSSWGQCNGPCIGPHLAVQHRQVFCQTRDGITLPSEQCSALPRPVSTQNCWSE 1647

Qy       1596 ACDVCWHTGPWKPCTAASGR GPQSPKVDCTHTRSCKTVAKRHPCVQKKKPTSWRHC LGP 1653
              || |  |      |    ||: ||    |||||:|:|:|  |:  |  |   :    |          :'  :::  |  :
Db       1648 ACSVHWKVSLWTDCTATCGNYGPQSRPVECVHAPINKAVPEHLCSWGPRPANWQRCNITP 1707

Qy       1654 SCDRDCTDTTHYCMEVKHLNEGSLDRYKQRCGQSC 1688
                  :  :|  |||  ||    || |  || |  ::|  |||  :|
Db       1708 CENMECRDTTRYCEKVKQLKLLQLSQFKSRCCGTC 1742
```

RESULT 8
US-09-764-853-554
; Sequence 554, Application US/09764853
; Patent No. US20020006722A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PZ006
; CURRENT APPLICATION NUMBER: US/09/764,853
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 939
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 554
; LENGTH: 1766
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (533)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L amino acids
US-09-764-853-554

```
  Query Match             41.3%;  Score 3860.5;  DB 10;  Length 1766;
  Best Local Similarity   42.1%;  Pred. No. 1.4e-211;
  Matches  765;  Conservative  266;  Mismatches  579;  Indels  207;  Gaps   33;

Qy         25 TAEKSPGAYPLPEPALSPQGSF  ELEPTIGEQPLTYRYDDQISRNIRSDEDKIGNWDAW 81
              |||:     || : :          |             | |     |    :      :|,    ||:||:|| ||||
Db          1 TAEEHLAAYSVSDSDSGKDPSMECCRRAIPGTLLLFLAFLLLSSRTARSEEDREGLWDAW 60

Qy         82 GDWSDCSRTCGGGASYSLRRCLTGRNCEGQNIRYKTCSNHDCPPDAEDFRAQQCSAYNDV 141
              |  ||:|||||||||||||||| :  ::|||:|||:||||  ||||:| |||||  |||:|||
Db         61 GPWSECSRTCGGGASYSLRRCLSSKSCEGRNIRYRTCSNVDCPPEAGDFRAQQCSAHNDV 120

Qy        142 QYQGHYYEWLPRYNDPAAPCALKCHAQGQNLVVELAPKVLDGTRCNTDSLDMCISGICQA 201
              :: |  :|||||   |||   ,|:||| |:|   ||||||||||||||||  |:||||||||:||
Db        121 KHHGQFYEWLPVSNDPDNPCSLKCQAKGTTLVVELAPKVLDGTRCYTESLDMCISGLCQI 180

Qy        202 VGCDRQLGSNAKEDNCGVCAGDGSTCRLVRGQSKSHVSPEKREENVIAVPLGSRSVRITV 261
              |||| ||||   |||||||| |||||||||||| :|  :'  |  ::  |:|:| ||| :|: :
Db        181 VGCDHQLGSTVKEDNCGVCNGDGSTCRLVRGQYKSQLSATKSDDTVVAIPYGSRHIRLVL 240

Qy        262 KGPAHLFIESKTLQGSKGEHSFNSPGVFVVENTTVEFQRGSERQTFKIPGPLMADFIFKT 321
              ||| ||::|:||||||:||||:|  :|  :  ':::|::|:|:      :::   ::  |||  '|||  |
Db        241 KGPDHLYLETKTLQGTKGENSLSSTGTFLVDNSSVDFQKFPDKEILRMAGPLTADFIVKI 300

Qy        322 RYTAAKDSVVQFFFYQPISHQWRQTDFFPCTVTCGGGYQLNSAECVDIRLKRVVPDHYCH 381
              |  :  :  ||  |||  '||||  |:|||:||||||||:  ||||||||  |:|   ':::  ||    |  :||
Db        301 RNSGSADSTVQFIFYQPITHRWRETDFFPCSATCGGGYQLTSAECYDLRSNRVVADQYCH 360

Qy        382 YYPENVKPKPKLKECSMDPCPSSDGFREIMPYDHFQPLPRWEHNPWTACSVSCGGGIQRR 441
              |''|||:|||||,|:::  ::;|||:                                   |:|  ,|||'||  ||:  ||||   ,
Db        361 YYPENIKPKPKLQECNLDPCPA                         RWEATPWTACSSSCGGGIQSR 403

Qy        442 SFVCVEESMHGEIPQVEEWKCMYAPKPKVMQTCNLFDCPKWIAMEWSQCTVTCGRGLRYR 501
              :   ||||  :  |  :     ||||||:|,|  ||      :  |  ||:|;|  ||:| |||,   |||'||:||||
Db        404 AVSCVEEDIQGHVTSVEEWKCMYTPKMPIAQPCNIFDCPKWLAQEWSPCTVTCGQGLRYR 463

Qy        502 VVLCINHRGEHVGGCNPQLKLHIKEECVIPIPCYKPKEKSPVEAKLPWLKQAQELEETRI 561
              |||||:|||  |  ||||:|:  |  ||||||||::|  |||||||||  ||||(|||  ||||||||
Db        464 VVLCIDHRGMHTGGCSPKTKPHIKEECIVPIPCYKPKEKLPVEAKLPWFKQAQELEEGAA 523

Qy        562 ATEEPTFTPEPWSACSTTCGPGVQVREVKCRVLLTFTQTETELPEEECEGPKLPTERPCL 621
                :|||:|||  ||||: |||  ,  |||  |:|:|||'|:|:|:    :||  :||||||    ::|
Db        524 VSEEPSFTPXAWSACTVPCGVGTQVRTVRCQVLLSFSQSVADLPIDECEGPKLASQR    580

Qy        622 LEACDESPASRELDTPLP              EPSETTYDWEYAGLTPCTATCLGGHQEATAVC 671
                |||    |  |  |  |:     |               :|  :    |||||  |||   |:   :|  ||  |||:      :
Db        581   ACYAGPCSGEIPEENPDETDGLEGGLQDFDELYDWEYEGFTKCSESCGGGVQEAVVSC 638

Qy        672 LHIQTQQTVNDSLCDMVHRPPAMSQACNTEPCPPRWHVGSWGPCSATCGVGIQTRDVYCL 731
              |: :  |||::       :  :||     ||:    :  :  :|  ||||   |||   :|     |    |||    ||||||:||||||:|
Db        639 LNKQTREPAEENLCVTSRRPPQLLKSCNLDPCPARWEIGKWGPCSLTCGVGLQTRDVFCS 698

Qy        732 H          EGETPAPPEE CRGEKPHALQACNQFTCPPGWHIEWQQCSRTCGGGTQNRRVT 785
                      ||           :|   |,   ||        :|'||':|:'||   |:       :||   |,||'|''   |   |   |
Db        699 HLLSREMNETVTLADELCRQPKPSTVQACNRFNCPPAWYPAQWQPCSRTCGGGVQKREVT 758
```

```
Qy        786 CRQLLIDGGFLNLSDHLQQGFKASSHKSCARTDCPFHLAVCDWSRCSVSCGVGIQRREQV 845
              |:|  :  |||||  |  :    |    |  :   ::|  :  |||        :  ||::||  |||  |  |  |    :
Db        759 CKQRMALCSFIFIPFTFTSASKFATQLACFFDFTPSFWIIFIWIEISIFCGEGIQIKSAI 818

Qy        846 CQRLAAFCRRIPLSEMMCRELPGFPLVRSCQMPECSSKSEMFIPICEQCPQIILSVQRVY 905
              |:::     |       ::    :|    ||        :|  |  :   |::              ||         |  |  :  ::||
Db        819 CRKMLKICLSTVVNSILCPPLPFSSSIRPCMLATCAR-PCRPSTK---HSPHIAAARKVY 874

Qy        906 IQTRFFKRINLTICSFAYLLPNTSVIIKCPVRRFQKSLIQWEKDGRCLQNSKRLGITKSC 965
              ||||  :::::    :|    |||||  |:|:::||  ||  :|  ||  |||||:  |  :|    :  :    |
Db        875 IQTRRQRFIHFVVSCFAYLLPFIAVVLRCFARRVRRFLIIWERDGQHLISSTHVTVAPFC 934

Qy        966 SIKIHGLAAPDIGVYRCIAGSAQETVVLKLIGTDNRLIARP----------ALREPMREY 1015
               ||||  |     |  |||  |  ||  |:|   |:||||  .   .|.|||          |  |.     .
Db        935 YLKIHRLKFSDAGVYTCSAGFARRHFVIKLIGGNRFIVARPLSPRSEEEVLAGRKGFRF 994

Qy       1016 PGMDHSEANSLGVTWHK--MRQMWNNKNDLYLDDDHISNQPFLRALLGHCSNSAGSTNSW 1073
                     |    |  :        |     |  :   |       |    ||  :|          ||               ||
Db        995 ALQIHRHQNSIFENISFALERGLAANGSSKY  DLLVSR       LLEQGGWFGELLASW 1046

Qy       1074 ELKNKQFEAAVKQ   CA--------YSMPTAQ--FDELIRNMSQLMETGEVSDDLASQL 1120
              |  ::          :    ||           ::|  |  |     |:::  |:||   |    |:  |   :   |
Db       1047 EAQRSAPRNTTSPPLSSAPQVIIHIPFTMVTEQFFILPIICNLSQQPE -ELRDLYSKHL 1104

Qy       1121 IYQLVAPLAKAQPTHMQ   WRGIQEEIPP---------   ---AAQLR----CETCSVSQSS 1162
              :  ||   |:  ::      |          :     :    |  |                ::  ||        |:  |    |:
Db       1105 VAQLAQEIFRSHLEHQDTILKPSERRTSPVTLSPHKHVSGFSSSLRTSSTGDAGGGSPRP 1164

Qy       1163 HAKNSGKITFKPKGPVLMRQ-SQPPSISFNKTINSRIGNTVYITKRTEVINILCDLITPS 1221
              |  |                    |  ::|:  |       :|  ::  :  :  :|  ||  :       |    :  :  |:  |
Db       1165 HRK------------PTILRKISAAQQISASEVV-THIGQTVAIASQILSVLLHCEAIGHF 1212

Qy       1222 EATYIWTKDGTILQPSVKIIILDGTGKIQIQNPTRKEQGIYECSVANHLGSDVESSSVLYA 1281
                 |  :|  ::|     :|  |  :|:|          :||  |      :  |  |  |:    |  ||  |    |  :|   |
Db       1213 RPTISWARNGEEVQFSDRIIIQPDDSIQIIAPVRADVGFYTCNATNALGYDSVSIAVTLA 1272

Qy       1282 EAPVILSVERNITKPFHNHISVVVGGIVEAALGANVTIRCPVKGVPQPNITWLKRGGSLS 1341
               |::  :       :     |      ::|  :|    ::        |  ||||  |  |  |||:     :||  :          |
Db       1273 GKPLVKTSRMTVINTEKPAVTVDIGSTIKTVQGVNVTINCQVAGVPEAFVTWFRNKSKL  1331

Qy       1342 GNVSLLFNGSLLLQNVSLENEGTYVCIATNALGKAVATSVLHLLERRWPESRIVFLQGHK 1401
              |:      |      |||||  |||       ::|  |    |    |    |    |          |:           ::  |    :|:              :::     ::
Db       1332 GSPHHLHRGSLLLINVSSSDQGIYSCRAANLHGPLTESTQSILIPPQVPTQLEDIRA- 1389

Qy       1402 KYILQATNTRTNS                   NRP                TCPPPP            1423
                :|  ||          |                   ||                 |   |   |
Db       1390 --LLAAFGPNLPSVLISPLGTQLVLDPGNSALIGCPIKGHPVPNITWFHGGQPIVTATGL 1447

Qy       1424 -------   ---------------                ---           QHPFWEPGNWSHCSATC 1440
                                                                     |:  :|            :  ||||:|
Db       1448 THHILAAGQILQVANLSGGSQGEFSCLAQNEAGVLMQKASLVIQDYWWSVDRLATCSASC 1507

Qy       1441 GHLGARIQRPQCVMANGQEVSEALCDHLQKPLAGFEPCNIRDCPARWFTSVWSQCSVSCG 1500
              |:  |  :    |  :|::  |    ||:  |  |          :|           ||  ||||:||       :  ||  |:  |||
Db       1508 GNRGVQQPRLRCLL NSTEVNPAHCAGKVRPAVQPIACNRRDCPSRWMVTSWSACTRSCG 1566

Qy       1501 EGYHSRQVTCKRTKANGTVQVVSPRACAP KDRPLGRKPCFGHPCVQW   EPGNRCPGRC 1557
                |      :|:||||::  |:  :|           ||         |              ||:          :  |            ||:|                 :|  |  |
Db       1567 GGVQTRRVTCQKLKASGISTPVSNDMCTQVAKRPVDTQACNQQLCVEWAFSSWGQCNGPC 1626

Qy       1558 MGRAVRMQQRHTACQHNGS   -  DSNCDDRKRPTLRRNCTSGACDVCWHTGPWKPCTAAC 1613
              :|    :  :|   |         ||                            |       ||       :||  |  ||  |  |          |     |||  |
Db       1627 IGPHLAVQHRQVFCQTRDGITLPSEQCSALPRPVSTQNCWSFACSVPWRVSIWTICTATC 1686

Qy       1614 GR GFQSRKVDCIHTRSCKPVAKRHCVQKKKPISWRHC LGPSCDRDCTDTTHYCMFVKH 1671
              |     ||||||:|:|:|  |:  |  |  |   :   |          :|  :|:  |  :  |    :   :|  |||  ||   ||
Db       1687 GNYGFQSRRVECVHARTNKAVPEHLCSWGPRPANWQRCNITPCLNMECRDTTRYCEKVKQ 1746

Qy       1672 LNLCSLDRYKQRCCQSC 1688
              |  ||  :  ::|  |||  :|
Db       1747 LKLCQLSQFKSRCCGTC 1763
```

```
RESULT 9
US-09-784-358-10
; Sequence 10, Application US/09784358
; Patent No. US20020099027A1
; GENERAL INFORMATION:
;  APPLICANT: Donoho, Gregory
;  APPLICANT:  Scoville, John
;  APPLICANT:  Turner, C. Alexander Jr.
;  APPLICANT:  Friedrich, Glenn
;  APPLICANT:  Zambrowicz, Brian
;  APPLICANT:  Sands, Arthur T.
;  TITLE OF INVENTION: NOVEL HUMAN THROMBOSPONDIN REPEAT PROTEINS AND
;  TITLE OF INVENTION:  POLYNUCLEOTIDES ENCODING THE SAME
;  FILE REFERENCE: LEX-0134-USA
;  CURRENT APPLICATION NUMBER: US/09/784,358
;  CURRENT FILING DATE:  2001-02-15
;  PRIOR APPLICATION NUMBER: US 60/183,282
;  PRIOR FILING DATE: 2000-02-17
;  NUMBER OF SEQ ID NOS: 17
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10
;   LENGTH: 650
;   TYPE: PRT
;   ORGANISM: homo sapiens
US-09-784-358-10

  Query Match             37.4%;  Score 3492;  DB 10;  Length 650;
  Best Local Similarity   89.5%;  Pred. No. 3.7e-191;
  Matches  632;  Conservative    0;  Mismatches    0;  Indels   74;  Gaps    1;

Qy          1 MASWTSPWWVLIGMVFMHSPLPQTTAEKSPGAYFLPEFALSPQGSFLEDTTGEQFLTYRY 60
              ||||||||||||||||||||||||||||||
Db          1 MASWTSPWWVLIGMVFMHSPLPQTTAEKSP------------------------------ 30

Qy         61 DDQTSRNTRSDEDKDGNWDAWGDWSDCSRTCGGGASYSLRRCLTGRNCEGQNIRYKTCSN 120
                                                          ||||||||||||||||
Db         31 --------------------------------------------GRNCEGQNIRYKTCSN 46

Qy        121 HDCPPDAEDFRAQQCSAYNDVQYQGHYYEWLPRYNDPAAPCALKCHAQGQNLVVELAPKV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         47 HDCPPDAEDFRAQQCSAYNDVQYQGHYYEWLPRYNDPAAPCALKCHAQGQNLVVELAPKV 106

Qy        181 LDGTRCNTDSLDMCISGICQAVGCDRQLGSNAKEDNCGVCAGDGSTCRLVRGQSKSHVSP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 LDGTRCNTDSLDMCISGICQAVGCDRQLGSNAKEDNCGVCAGDGSTCRLVRGQSKSHVSP 166

Qy        241 EKREENVIAVPLGSRSVRITVKGPAHLFIESKTLQGSKGEHSFNSPGVFVVENTTVEFQR 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 EKREENVIAVPLGSRSVRITVKGPAHLFIESKTLQGSKGEHSFNSPGVFVVENTTVEFQR 226

Qy        301 GSERQTFKIPGPLMADFIFKTRYTAAKDSVVQFFFYQPISHQWRQTDFFPCTVTCGGGYQ 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 GSERQTFKIPGPLMADFIFKTRYTAAKDSVVQFFFYQPISHQWRQTDFFPCTVTCGGGYQ 286

Qy        361 LNSAECVDIRLKRVVPDHYCHYYPENVKPKPKLKECSMDPCPSSDGFKEIMPYDHFQPLP 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 LNSAECVDIRLKRVVPDHYCHYYPENVKPKPKLKECSMDPCPSSDGFKEIMPYDHFQPLP 346

Qy        421 RWEHNPWTACSVSCGGGIQRRSFVCVEESMHGEILQVEEWKCMYAPKPKVMQTCNLFDCP 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        347 RWEHNPWTACSVSCGGGIQRRSFVCVEESMHGEILQVEEWKCMYAPKPKVMQTCNLFDCP 406

Qy        481 KWIAMEWSQCTVTCGRGLRYRVVLCINHRGEHVGGCNPQLKLHIKEECVIPIPCYKPKEK 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        407 KWIAMEWSQCTVTCGRGLRYRVVLCINHRGEHVGGCNPQLKLHIKEECVIPIPCYKPKEK 466

Qy        541 SPVEAKLPWLKQAQELEETRIATEEPTFIPEPWSACSTTCGPGVQVREVKCRVLLTFTQT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        467 SPVEAKLPWLKQAQELEETRIATEEPTFIPEPWSACSTTCGPGVQVREVKCRVLLTFTQT 526

Qy        601 ETELPEEECEGPKLPTERPCLLEACDESPASRELDIPLPEDSETTYDWEYAGFTPCTATC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        527 EIELPEEECEGPKLPIERPCLLEACDESPASREELDIPLPEDSEIIYDWEYAGFIPCIAIC 586

Qy        661 LGGHQEATAVCLHIQIQQIVNDSLCDMVHRPPAMSQACNIEPCPPR 706
              |:||||||||||:|||  ....|||||||||||||:||||| ||||
Db        587 LGGHQEATAVCLHIQIQQIVNDSLCDMVHRPPAMSQACNIEPCPPR 632


RESULT 10
US-09-796-753-20
; Sequence 20, Application US/09796753
; Publication No. US20030027998A1
; GENERAL INFORMATION:
;  APPLICANT: McCarthy, Sean A.
;  TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
;  FILE REFERENCE: 7853-227-999
;  CURRENT APPLICATION NUMBER: US/09/796,753
;  CURRENT FILING DATE: 2001-03-01
;  PRIOR APPLICATION NUMBER: 09/183,175
;  PRIOR FILING DATE: 1998-10-30
;  PRIOR APPLICATION NUMBER: 09/223,004
;  PRIOR FILING DATE: 1998-12-30
;  PRIOR APPLICATION NUMBER: 09/223,546
;  PRIOR FILING DATE: 1998-12-30
;  PRIOR APPLICATION NUMBER: 09/224,219
;  PRIOR FILING DATE: 1998-12-30
;  PRIOR APPLICATION NUMBER: 09/250,386
;  PRIOR FILING DATE: 1999-02-26
;  PRIOR APPLICATION NUMBER: 09/261,[illegible]
;  PRIOR FILING DATE: 1999-03-01
;  PRIOR APPLICATION NUMBER: 09/312,959
;  PRIOR FILING DATE: 1999-05-14
;  PRIOR APPLICATION NUMBER: 09/335,[illegible]
;  PRIOR FILING DATE: 1999-06-18
;  PRIOR APPLICATION NUMBER: 09/342,687
;  PRIOR FILING DATE: 1999-06-29
;  PRIOR APPLICATION NUMBER: 09/345,191
;  PRIOR FILING DATE: 1999-06-30
;  PRIOR APPLICATION NUMBER: 09/365,234
;  PRIOR FILING DATE: 1999-07-30
;  PRIOR APPLICATION NUMBER: 09/399,273
;  PRIOR FILING DATE: 1999-09-21
;  PRIOR APPLICATION NUMBER: 09/410,[illegible]
;  PRIOR FILING DATE: 1999-09-30
;  PRIOR APPLICATION NUMBER: 09/471,179
;  PRIOR FILING DATE: 1999-12-23
;  PRIOR APPLICATION NUMBER: 09/471,673
;  PRIOR FILING DATE: 1999-12-29
;  PRIOR APPLICATION NUMBER: 09/471,872
;  PRIOR FILING DATE: 1999-12-29
;  PRIOR APPLICATION NUMBER: 09/511,963
;  PRIOR FILING DATE: 2000-02-25
;  PRIOR APPLICATION NUMBER: 09/516,745
;  PRIOR FILING DATE: 2000-03-01
;  PRIOR APPLICATION NUMBER: 09/572,[illegible]
;  PRIOR FILING DATE: 2000-05-14
;  PRIOR APPLICATION NUMBER: 09/597,993
;  PRIOR FILING DATE: 2000-06-19
;  PRIOR APPLICATION NUMBER: 09/599,[illegible]
;  PRIOR FILING DATE: 2000-06-22
;  PRIOR APPLICATION NUMBER: 09/632,334
;  PRIOR FILING DATE: 2000-07-31
;  PRIOR APPLICATION NUMBER: 09/638,565
;  PRIOR FILING DATE: 2000-08-29
;  PRIOR APPLICATION NUMBER: 09/656,317
;  PRIOR FILING DATE: 2000-06-29
;  PRIOR APPLICATION NUMBER: 09/665,666
;  PRIOR FILING DATE: 2000-09-20
;  PRIOR APPLICATION NUMBER: 09/677,251
;  PRIOR FILING DATE: 2000-09-30
;  NUMBER OF SEQ ID NOS: 162
; SEQ ID NO 20
;   LENGTH: 874
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-796-753-20

  Query Match             30.9%;  Score 2891;  DB 9;  Length 874;
  Best Local Similarity   58.4%;  Pred. No. 8.6e-157;
  Matches  493;  Conservative  126;  Mismatches  195;  Indels   30;  Gaps    5;

Qy         64 LGRNTKSGGDKIGNWLAWGLWSLCSKTCGGGASYSLRRCLIGRNCEGQNIRYKTCSNHDC 123
              :|:  |:|:|:| ||:|: : :|||||||||||||||||: ::|||:|||| ||| :|
Db         22 SSRTARSEEDRDGLWDAWGPWSECSRTCGGGASYSLRRCLSSKSCEGRNIRYRTCSNVDC 81

Qy        124 PPDAEDFRAQQCSAYNDVQYQGHYYEWLPRYNDPAAPCALKCHAQGQNLVVELAPKVLDG 183
              ||:| ||||||||:|||:: :||||   |||  ||:||:|:  |||||||||||
Db         82 PPEAGDFRAQQCSAHNDVKHHGQFYEWLPVSNDPDNPCSLKCQAKGTTLVVELAPKVLDG 141

Qy        184 TRCNTDSLDMCISGICQAVGCDRQLGSNAKEDNCGVCAGDGSTCRLVRGQSKSHVSPEKR 243
              |||  |:|||||||||:|| |||| ||||    ||||||||: ||||||||||| |: :|  |
Db        142 TRCYTESLDMCISGLCQIVGCDHQLGSTVKEDNCGVCNGDGSTCRLVRGQYKSQLSATKS 201

Qy        244 EENVIAVPLGSRSVRITVKGPAHLFIESKTLQGSKGEHSFNSPGVFVVENTTVEFQRGSE 303
              :: |:|:|  ||: :: :|| |::|:||:|||:|||:|:  :   :: :|::|:|||:  :
Db        202 DDTVVAIPYGSRHIRLVLKGPDHLYLETKTLQGTKGENSLSSTGTFLVDNSSVDFQKFPD 261

Qy        [illegible] 
              :: ::  ||| |  |:   : :  |: |||  ||||: :::|:|||:||: |||  ||||  |
Db        262 KEILRMAGPLTADFIVKIRNSGSADSTVQFIFYQPIIHRWRETDFFPCSATCGGGYQLTS 321

Qy        [illegible] TPCIMPYD 414
              ||| |:|  ||| | |||||||||:||||:||::|||              |:|||:::||||||
Db        322 AECYDLRSNRVVADQYCHYYPENIKPKPKLQECNLDPCPASDGYKQIMPYD 381

Qy        [illegible] 474
              : ||||||  ||||||| |||||||| |:   ||||  : |  :  ||||||||  ||   : | |
Db        382 LYHPLPRWEATPWTACSSSCGGGIQSRAVSCVEEDIQGHVTSVEEWKCMYTPKMPIAQPC 441

Qy        475 NLFDCPKWIAMEWSQCTVTCGQGLRYRVVLCINHRGEHVGGCNPQLKLHIKEECVIPIPC 534
              |:|||||||:| |||  |||||||:||||||||||||||:||| |  |::|:  |  ||||||::|| ||
Db        442 NIFDCPKWLAQEWSPCTVTCGQGLRYRVVLCIDHRGMHTGGCSPKTKPHIKEECIVPTPC 501

Qy        535 YKPKEKSPVEAKLPWLKQAQELEETRIATEEPTFIPEPWSACSTTCGPGVQVREVKCRVL 594
              |||||| |||||::|: ||||:|:|          :|:|::||| ||||: |||  |  ||: |:|:||
Db        502 YKPKEKLPVEAKLPWFKQAQELEEGAAVSEEPSFIPEAWSACTVTCGVGTQVRIVRCQVL 561

Qy        595 LTFTQTETELPEEECEGPKLPTERPCLLEACDESPASRELDIPLP           EDSET 644
              :|:|:  :||  :|:||||  ::|        ||    | | |:    |               :| :
Db        562 LSFSQSVADLPIDESEGPKEAGQK     ACYAGPCSGEIPEENEDEEDGLEPGLQDEDF 616

Qy        645 IYDWEYAGFTPCTATCAGGHQEATAVCLHIQTQQTVNDSLCDMVHRPPAMSQACNTEPCP 704
              ||||| |||  |: :|  ||  |||:  ||: |||::      ::||      |||  :  ::|| :|||
Db        617 LYDWKTLSLKNGAKQGKRVQEAVVSLNRQIEDAPENSVIGRRPPQLLKSNLDPCP 676

Qy        705 PRWHVGSWGPCSATCGVGIQTRDVYCLH           PGETPAPPEPCRDEKPHALQACNQFD 758
              ||  :|  | | :| ||||||:|||| :|  :            ||       :   ||   ||  :|||:|:
Db        677 ARWEIGEWGECLSVVGQIRLVHCHLLDGLMREVPIADELCPQPKPSIVQACNRRN 736

Qy        759 CPPGWHIEEWQQCSRTCGGGTQNRRVTCRQLLTDGSFLNLSDELCQGPKASSHKSCARTD 818
              |:|  |:  :||  |:|:||||: | |  | |:| : :||||  :     :       : :|  :  |
Db        737 [illegible]AWYIAQW[illegible]AVQRRAV[illegible]RQPHAC[illegible]ACRIA[illegible]TRKDC 796

Qy        819 CPPHLAVGDWSKCSVSCGVGIQRRKQVCQRLAASGRRIPLSEMMCRDLPGLPLVRSCQMP 878
              ||       :  |||::|||  |||  ||  |  |  :||:::      |       ::  :|     ||         :|  |  |  :
Db        797 CESEWLLSDWIECSTSCGEGTQTRSAICRKMLKTGTSTVVNSTLCPPLPFSSSIRPCMLA 856

Qy        879 ECSK 882
               |::
Db        857 TCAR 860


RESULT 11
US-09-784-358-4
; Sequence 4, Application US/09784358
; Patent No. US20020099027A1
; GENERAL INFORMATION:
```

```
;  APPLICANT: Donoho, Gregory
;  APPLICANT:  Scoville, John
;  APPLICANT:  Turner, C. Alexander Jr.
;  APPLICANT:  Friedrich, Glenn
;  APPLICANT:  Zambrowicz, Brian
;  APPLICANT:  Sands, Arthur T.
;  TITLE OF INVENTION: NOVEL HUMAN THROMBOSPONDIN REPEAT PROTEINS AND
;  TITLE OF INVENTION:  POLYNUCLEOTIDES ENCODING THE SAME
;  FILE REFERENCE: LEX-0134-USA
;  CURRENT APPLICATION NUMBER: US/09/784,358
;  CURRENT FILING DATE:  2001-02-15
;  PRIOR APPLICATION NUMBER: US 60/183,282
;  PRIOR FILING DATE: 2000-02-17
;  NUMBER OF SEQ ID NOS: 17
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
;   LENGTH: 446
;   TYPE: PRT
;   ORGANISM: homo sapiens
US-09-784-358-4

  Query Match             25.1%;  Score 2344;  DB 10;  Length 446;
  Best Local Similarity   100.0%;  Pred. No. 5e-126;
  Matches  421;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MASWTSPWWVLIGMVFMHSPLPQTIAEKSPGAYFLPEFALSPQGSFLEDTTGEQFLTYRY 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MASWTSPWWVLIGMVFMHSPLPQTIAEKSPGAYFLPEFALSPQGSFLEDTTGEQFLTYRY 60

Qy         61 DDQTSRNTRSDEDKDGNWDAWGDWSDCSRTCGGGASYSLRRCLTGRNCEGQNIRYKTCSN 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DDQTSRNTRSDEDKDGNWDAWGDWSDCSRTCGGGASYSLRRCLTGRNCEGQNIRYKTCSN 120

Qy        121 HDCPPDAEDFRAQQCSAYNDVQYQGHYYEWLPRYNDPAAPCALKCHAQGQNLVVELAPKV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 HDCPPDAEDFRAQQCSAYNDVQYQGHYYEWLPRYNDPAAPCALKCHAQGQNLVVELAPKV 180

Qy        181 LDGTRCNTDSLDMCISGICQAVGCDRQLGSNAKEDNCGVCAGDGSTCRLVRGQSKSHVSP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 LDGTRCNTDSLDMCISGICQAVGCDRQLGSNAKEDNCGVCAGDGSTCRLVRGQSKSHVSP 240

Qy        241 EEREENVIAVPLGSRSVRITVKGPAHLFIESKTLQGSKGEHSFNSPGVFVVENTTVEFQR 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 EEREENVIAVPLGSRSVRITVKGPAHLFIESKTLQGSKGEHSFNSPGVFVVENTTVEFQR 300

Qy        301 GSEPLTFEIPGHLMALIIRKTRYIAARDSVVQEFFYQPISHQWRQIDFFPCIVICGGGYQ 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GSEPLTFEIPGHLMALIIRKTRYIAARDSVVQEFFYQPISHQWRQIDFFPCIVICGGGYQ 360

Qy        361 LNSAECVDIFLKEVVIDHYTCHYYFENVEIRKNLRKLSMDEIPGSQGIKEIMPYDHFQPLP 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LNSAECVDIFLKEVVIDHYTCHYYFENVEIRKNLRKLSMDEIPGSQGIKEIMPYDHFQPLP 420

Qy        421 R 421
              |
Db        421 R 421


RESULT 12
US-09-978-295A-301
; Sequence 301, Application US/09978295A
; Patent No. US20020156006A1
; GENERAL INFORMATION:
;   APPLICANT: Ashkenazi, Avi
;   APPLICANT:  Baker Kevin P.
;   APPLICANT:  Botstein, David
;   APPLICANT:  Desnoyers, Luc
;   APPLICANT:  Eaton, Dan
;   APPLICANT:  Ferrara, Napoleon
;   APPLICANT:  Filvaroff, Ellen
;   APPLICANT:  Fong, Sherman
;   APPLICANT:  Gao, Wei-Qiang
;   APPLICANT:  Gerber, Hanspeter
;   APPLICANT:  Gerritsen, Mary E.
;   APPLICANT:  Goddard, Audrey
;   APPLICANT:  Godowski, Paul J.
;   APPLICANT:  Grimaldi, J. Christopher
;   APPLICANT:  Gurney, Austin L.
;   APPLICANT:  Hillan, Kenneth J
;   APPLICANT:  Kljavin, Ivar J.
;   APPLICANT:  Kuo, Sophia S.
;   APPLICANT:  Napier, Mary A.
;   APPLICANT:  Pan, James;
;   APPLICANT:  Paoni, Nicholas F.
;   APPLICANT:  Roy, Margaret Ann
;   APPLICANT:  Shelton, David L.
;   APPLICANT:  Stewart, Timothy A.
;   APPLICANT:  Tumas, Daniel
;   APPLICANT:  Williams, P. Mickey
;   APPLICANT:  Wood, William I.
;   TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;   TITLE OF INVENTION:  Acids Encoding the Same
;   FILE REFERENCE: P2630P1C11
;   CURRENT APPLICATION NUMBER: US/09/978,295A
;   CURRENT FILING DATE:  2001-10-15
;   PRIOR APPLICATION NUMBER: 09/918585
;   PRIOR FILING DATE: 2001-07-30
;   PRIOR APPLICATION NUMBER: 60/062250
;   PRIOR FILING DATE: 1997-10-17
;   PRIOR APPLICATION NUMBER: 60/064249
;   PRIOR FILING DATE: 1997-11-03
;   PRIOR APPLICATION NUMBER: 60/065311
;   PRIOR FILING DATE: 1997-11-13
;   PRIOR APPLICATION NUMBER: 60/066364
;   PRIOR FILING DATE: 1997-11-21
;   PRIOR APPLICATION NUMBER: 60/077450
;   PRIOR FILING DATE: 1998-03-10
;   PRIOR APPLICATION NUMBER: 60/077632
;   PRIOR FILING DATE: 1998-03-11
;   PRIOR APPLICATION NUMBER: 60/077641
;   PRIOR FILING DATE: 1998-03-11
;   PRIOR APPLICATION NUMBER: 60/077649
;   PRIOR FILING DATE: 1998-03-11
;   PRIOR APPLICATION NUMBER: 60/077791
;   PRIOR FILING DATE: 1998-03-12
;   PRIOR APPLICATION NUMBER: 60/078004
;   PRIOR FILING DATE: 1998-03-13
;   PRIOR APPLICATION NUMBER: 60/078886
;   PRIOR FILING DATE: 1998-03-20
;   PRIOR APPLICATION NUMBER: 60/078936
;   PRIOR FILING DATE: 1998-03-20
;   PRIOR APPLICATION NUMBER: 60/078910
;   PRIOR FILING DATE: 1998-03-20
;   PRIOR APPLICATION NUMBER: 60/078939
;   PRIOR FILING DATE: 1998-03-20
;   PRIOR APPLICATION NUMBER: 60/079294
;   PRIOR FILING DATE: 1998-03-25
;   PRIOR APPLICATION NUMBER: 60/079656
;   PRIOR FILING DATE: 1998-03-26
;   PRIOR APPLICATION NUMBER: 60/079664
;   PRIOR FILING DATE: 1998-03-27
;   PRIOR APPLICATION NUMBER: 60/079689
;   PRIOR FILING DATE: 1998-03-27
;   PRIOR APPLICATION NUMBER: 60/079663
;   PRIOR FILING DATE: 1998-03-27
;   PRIOR APPLICATION NUMBER: 60/079728
;   PRIOR FILING DATE: 1998-03-27
;   PRIOR APPLICATION NUMBER: 60/079786
;   PRIOR FILING DATE: 1998-03-27
;   PRIOR APPLICATION NUMBER: 60/079920
;   PRIOR FILING DATE: 1998-03-30
;   PRIOR APPLICATION NUMBER: 60/079923
;   PRIOR FILING DATE: 1998-03-30
;   PRIOR APPLICATION NUMBER: 60/080105
;   PRIOR FILING DATE: 1998-03-31
;   PRIOR APPLICATION NUMBER: 60/080107
;  PRIOR FILING DATE: 1998-03-31
;  PRIOR APPLICATION NUMBER: 60/080165
;  PRIOR FILING DATE: 1998-03-31
;  PRIOR APPLICATION NUMBER: 60/080194
;  PRIOR FILING DATE: 1998-03-31
;  PRIOR APPLICATION NUMBER: 60/080327
;  PRIOR FILING DATE: 1998-04-01
;  PRIOR APPLICATION NUMBER: 60/080328
;  PRIOR FILING DATE: 1998-04-01
;  PRIOR APPLICATION NUMBER: 60/080333
;  PRIOR FILING DATE: 1998-04-01
;  PRIOR APPLICATION NUMBER: 60/080334
;  PRIOR FILING DATE: 1998-04-01
;  PRIOR APPLICATION NUMBER: 60/081070
;  PRIOR FILING DATE: 1998-04-08
;  PRIOR APPLICATION NUMBER: 60/081049
;  PRIOR FILING DATE: 1998-04-08
;  PRIOR APPLICATION NUMBER: 60/081071
;  PRIOR FILING DATE: 1998-04-08
;  PRIOR APPLICATION NUMBER: 60/081195
;  PRIOR FILING DATE: 1998-04-08
;  PRIOR APPLICATION NUMBER: 60/081203
;  PRIOR FILING DATE: 1998-04-09
;  PRIOR APPLICATION NUMBER: 60/081229
;  PRIOR FILING DATE: 1998-04-09
;  PRIOR APPLICATION NUMBER: 60/081955
;  PRIOR FILING DATE: 1998-04-15
;  PRIOR APPLICATION NUMBER: 60/081817
;  PRIOR FILING DATE: 1998-04-15
;  PRIOR APPLICATION NUMBER: 60/081819
;  PRIOR FILING DATE: 1998-04-15
;  PRIOR APPLICATION NUMBER: 60/081952
;  PRIOR FILING DATE: 1998-04-15
;  PRIOR APPLICATION NUMBER: 60/081838
;  PRIOR FILING DATE: 1998-04-15
;  PRIOR APPLICATION NUMBER: 60/082568
;  PRIOR FILING DATE: 1998-04-21
;  PRIOR APPLICATION NUMBER: 60/082569
;  PRIOR FILING DATE: 1998-04-21
;  PRIOR APPLICATION NUMBER: 60/082704
;  PRIOR FILING DATE: 1998-04-22
;  PRIOR APPLICATION NUMBER: 60/082804
;  PRIOR FILING DATE: 1998-04-22
;  PRIOR APPLICATION NUMBER: 60/082700
;  PRIOR FILING DATE: 1998-04-22
;  PRIOR APPLICATION NUMBER: 60/082797
;  PRIOR FILING DATE: 1998-04-22
;  PRIOR APPLICATION NUMBER: 60/082796
;  PRIOR FILING DATE: 1998-04-23
;  PRIOR APPLICATION NUMBER: 60/083336
;  PRIOR FILING DATE: 1998-04-27
;  PRIOR APPLICATION NUMBER: 60/083322
;  PRIOR FILING DATE: 1998-04-28
;  PRIOR APPLICATION NUMBER: 60/083392
;  PRIOR FILING DATE: 1998-04-29
;  PRIOR APPLICATION NUMBER: 60/083495
;  PRIOR FILING DATE: 1998-04-29
;  PRIOR APPLICATION NUMBER: 60/083496
;  PRIOR FILING DATE: 1998-04-29
;  PRIOR APPLICATION NUMBER: 60/083499
;  PRIOR FILING DATE: 1998-04-29
;  PRIOR APPLICATION NUMBER: 60/083545
;  PRIOR FILING DATE: 1998-04-29
;  PRIOR APPLICATION NUMBER: 60/083554
;  PRIOR FILING DATE: 1998-04-29
;  PRIOR APPLICATION NUMBER: 60/083558
;  PRIOR FILING DATE: 1998-04-29
;  PRIOR APPLICATION NUMBER: 60/083559
;  PRIOR FILING DATE: 1998-04-29
;  PRIOR APPLICATION NUMBER: 60/083500
;  PRIOR FILING DATE: 1998-04-29
;  PRIOR APPLICATION NUMBER: 60/083742
;  PRIOR FILING DATE: 1998-04-30
```

```
;  PRIOR APPLICATION NUMBER: 60/084366
;  PRIOR FILING DATE: 1998-05-05
;  PRIOR APPLICATION NUMBER: 60/084414
;  PRIOR FILING DATE: 1998-05-06
;  PRIOR APPLICATION NUMBER: 60/084441
;  PRIOR FILING DATE: 1998-05-06
;  PRIOR APPLICATION NUMBER: 60/084637
;  PRIOR FILING DATE: 1998-05-07
;  PRIOR APPLICATION NUMBER: 60/084639
;  PRIOR FILING DATE: 1998-05-07
;  PRIOR APPLICATION NUMBER: 60/084640
;  PRIOR FILING DATE: 1998-05-07
;  PRIOR APPLICATION NUMBER: 60/084598
;  PRIOR FILING DATE: 1998-05-07
;  PRIOR APPLICATION NUMBER: 60/084600
;  PRIOR FILING DATE: 1998-05-07
;  PRIOR APPLICATION NUMBER: 60/084627
;  PRIOR FILING DATE: 1998-05-07
;  PRIOR APPLICATION NUMBER: 60/084643
;  PRIOR FILING DATE: 1998-05-07
;  PRIOR APPLICATION NUMBER: 60/085339
;  PRIOR FILING DATE: 1998-05-13
;  PRIOR APPLICATION NUMBER: 60/085338
;  PRIOR FILING DATE: 1998-05-13
;  PRIOR APPLICATION NUMBER: 60/085323
;  PRIOR FILING DATE: 1998-05-13
;  PRIOR APPLICATION NUMBER: 60/085582
;  PRIOR FILING DATE: 1998-05-15
;  PRIOR APPLICATION NUMBER: 60/085700
;  PRIOR FILING DATE: 1998-05-15
;  PRIOR APPLICATION NUMBER: 60/085689
;  PRIOR FILING DATE: 1998-05-15
;  PRIOR APPLICATION NUMBER: 60/085579
;  PRIOR FILING DATE: 1998-05-15
;  PRIOR APPLICATION NUMBER: 60/085580
;  PRIOR FILING DATE: 1998-05-15
;  PRIOR APPLICATION NUMBER: 60/085573
;  PRIOR FILING DATE: 1998-05-15
;  PRIOR APPLICATION NUMBER: 60/085704
;  PRIOR FILING DATE: 1998-05-15
;  PRIOR APPLICATION NUMBER: 60/085697

  Query Match             22.0%;  Score 2058;  DB 9;  Length 525;
  Best Local Similarity   68.7%;  Pred. No. 1.2e-109;
  Matches  346;  Conservative   71;  Mismatches   87;  Indels    0;  Gaps    0;

Qy         64 FSRNTRGDEDKDGNWDAWGDWSDCSRTCGGGASYSLRRCLTGKNCEGQNIRYKTCSNHDC 123
              :||  ||:||:|| |||| ||:|||||||||||||||||: :||:||||:|||| ||
Db         22 SSRTARSEEDRDGLWDAWGPWSECSRTCGGGASYSLRRCLSSKSCEGRNIRYRTCSNVDC 81

Qy        124 PPDAEDFRAQQCSAYNDVQYQGHYYEWLPRYNDPAAPCALKCHAQGQNLVVELAPKVLDG 183
              ||:| |||||||:||::| :|||||   |||   ||:||| |:|   ||||||||||||
Db         82 PPEAGDFRAQQCSAHNDVKHHGQFYEWLPVSNDPDNPCSLKCQAKGTTLVVELAPKVLDG 141

Qy        184 TRCNTDSLDMCISGICQAVGCDRQLGSNAKEDNCGVCAGDGSTCRLVRGQSKSHVSPEKR 243
              |||  |:||||||||:||  ||||  ||||   ||||||||  |||||  ||||||  |: :|   |
Db        142 TRCYTESLDMCISGLCQIVGCDHQLGSTVKEDNCGVCNGDGSTCRLVRGQYKSQLSATKS 201

Qy        244 EENVIAVPLGSRSVRITVKGPAHLFIESKTLQGSKGEHSFNSPGVFVVENTTVEFQRGSE 303
              :: |:|:| ||| :|: :||| ||::|:|||||:|||:| :|      |:|:|::::||:   :
Db        202 DDTVVALPYGSRHIRLVLKGPDHLYLETKTLQGTKGENSLSSTGTFLVDNSSVDFQKFPD 261

Qy        304 RQTFKIPGPLMADFIFKTRYTAAKDSVVQFFFYQPISHQWRQTDFFPCTVTCGGGYQLNS 363
              ::   :: ||| |||| | | : : || ||| ||||| |:||:| ||||: |||||||| |
Db        262 KEILRMAGPLTADFIVKIRNSGSADSTVQFIFYQPIIHRWRETDFFPCSATCGGGYQLTS 321

Qy        364 AECVDIRLKRVVPDHYCHYYPENVKPKPKLKECSMDPCPSDDGFKEIMPYDHFQPLPRWE 423
              ||| |:|  ||| | |||||||||:||||||:||::||||:|||.|:||||    .||||||
Db        322 AECYDLRSNRVVADQYCHYYPENIKPKPKLQECNLDPCPASDGYKQIMPYDLYHPLPRWE 381

Qy        424 HNPWTACSVSCGGGIQRRSFVCVEESMHGEILQVEEWKCMYAPKPKVMQTCNLFDCPKWI 483
                ||||||  |||||||| |:  ||||  : | :    |||||||||  |     : |  ||.||||||:
Db        382 ATPWTACSSSCGGGIQSRAVSCVEEDIQGHVTSVEEWKCMYTPKMPIAQPCNIFDCPKWL 441
```

```
Qy        484 AMEWSQCTVTCGRGLRYRVVLCINHRGEHVGGCNPQLKLHIKEECVIPIPCYKPKEKSPV 543
              | ||| ||||||:|||||||||||:||| | |||:|: | ||||||::|||||||||| ||
Db        442 AQEWSPCTVTCGQGLRYRVVLCIDHRGMHTGGCSPKTKPHIKEECIVPIPCYKPKEKLPV 501

Qy        544 EAKLPWLKQAQELEETRIATEEPT 567
              |||||| ||||||||    :|||:
Db        502 EAKLPWFKQAQELEEGAAVSEEPS 525


RESULT 13
US-09-978-697-301
; Sequence 301, Application US/09978697
; Patent No. US20020169284A1
; GENERAL INFORMATION:
;  APPLICANT: Ashkenazi, Avi
;  APPLICANT:  Baker Kevin P
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan
;  APPLICANT:  Ferrara, Napoleon
;  APPLICANT:  Filvaroff, Ellen
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Gerber, Hanspeter
;  APPLICANT:  Gerritsen, Mary E.
;  APPLICANT:  Goddard, Audrey
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, J. Christopher
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth J
;  APPLICANT:  Kljavin, Ivar J.
;  APPLICANT:  Kuo, Sophia S.
;  APPLICANT:  Napier, Mary A.
;  APPLICANT:  Pan, James;
;  APPLICANT:  Paoni, Nicholas F.
;  APPLICANT:  Roy, Margaret Ann
;  APPLICANT:  Shelton, David L.
;  APPLICANT:  Stewart, Timothy A.
;  APPLICANT:  Tumas, Daniel
;  APPLICANT:  Williams, P. Mickey
;  APPLICANT:  Wood, William I.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: P2630P1C27
;  CURRENT APPLICATION NUMBER: US/09/978,697
;  CURRENT FILING DATE:  2001-10-16
;  PRIOR APPLICATION NUMBER: 09/918585
;  PRIOR FILING DATE: 2001-07-30
;  PRIOR APPLICATION NUMBER: 60/062250
;  PRIOR FILING DATE: 1997-10-17
;  PRIOR APPLICATION NUMBER: 60/064249
;  PRIOR FILING DATE: 1997-11-03
;  PRIOR APPLICATION NUMBER: 60/065311
;  PRIOR FILING DATE: 1997-11-13
;  PRIOR APPLICATION NUMBER: 60/066364
;  PRIOR FILING DATE: 1997-11-21
;  PRIOR APPLICATION NUMBER: 60/077450
;  PRIOR FILING DATE: 1998-03-10
;  PRIOR APPLICATION NUMBER: 60/077632
;  PRIOR FILING DATE: 1998-03-11
;  PRIOR APPLICATION NUMBER: 60/077641
;  PRIOR FILING DATE: 1998-03-11
;  PRIOR APPLICATION NUMBER: 60/077649
;  PRIOR FILING DATE: 1998-03-11
;  PRIOR APPLICATION NUMBER: 60/077791
;  PRIOR FILING DATE: 1998-03-12
;  PRIOR APPLICATION NUMBER: 60/078004
;  PRIOR FILING DATE: 1998-03-13
;  PRIOR APPLICATION NUMBER: 60/078886
;  PRIOR FILING DATE: 1998-03-20
;  PRIOR APPLICATION NUMBER: 60/078936
;  PRIOR FILING DATE: 1998-03-20
;  PRIOR APPLICATION NUMBER: 60/078910
;  PRIOR FILING DATE: 1998-03-20
;  PRIOR APPLICATION NUMBER: 60/078939
;  PRIOR FILING DATE: 1998-03-20
;  PRIOR APPLICATION NUMBER: 60/079294
;  PRIOR FILING DATE: 1998-03-25
;  PRIOR APPLICATION NUMBER: 60/079656
;  PRIOR FILING DATE: 1998-03-26
;  PRIOR APPLICATION NUMBER: 60/079664
;  PRIOR FILING DATE: 1998-03-27
;  PRIOR APPLICATION NUMBER: 60/079689
;  PRIOR FILING DATE: 1998-03-27
;  PRIOR APPLICATION NUMBER: 60/079663
;  PRIOR FILING DATE: 1998-03-27
;  PRIOR APPLICATION NUMBER: 60/079728
;  PRIOR FILING DATE: 1998-03-27
;  PRIOR APPLICATION NUMBER: 60/079786
;  PRIOR FILING DATE: 1998-03-27
;  PRIOR APPLICATION NUMBER: 60/079920
;  PRIOR FILING DATE: 1998-03-30
;  PRIOR APPLICATION NUMBER: 60/079923
;  PRIOR FILING DATE: 1998-03-30
;  PRIOR APPLICATION NUMBER: 60/080105
;  PRIOR FILING DATE: 1998-03-31
;  PRIOR APPLICATION NUMBER: 60/080107
;  PRIOR FILING DATE: 1998-03-31
;  PRIOR APPLICATION NUMBER: 60/080165
;  PRIOR FILING DATE: 1998-03-31
;  PRIOR APPLICATION NUMBER: 60/080194
;  PRIOR FILING DATE: 1998-03-31
;  PRIOR APPLICATION NUMBER: 60/080327
;  PRIOR FILING DATE: 1998-04-01
;  PRIOR APPLICATION NUMBER: 60/080328
;  PRIOR FILING DATE: 1998-04-01
;  PRIOR APPLICATION NUMBER: 60/080333
;  PRIOR FILING DATE: 1998-04-01
;  PRIOR APPLICATION NUMBER: 60/080334
;  PRIOR FILING DATE: 1998-04-01
;  PRIOR APPLICATION NUMBER: 60/081070
;  PRIOR FILING DATE: 1998-04-08
;  PRIOR APPLICATION NUMBER: 60/081049
;  PRIOR FILING DATE: 1998-04-08
;  PRIOR APPLICATION NUMBER: 60/081071
;  PRIOR FILING DATE: 1998-04-08
;  PRIOR APPLICATION NUMBER: 60/081195
;  PRIOR FILING DATE: 1998-04-08
;  PRIOR APPLICATION NUMBER: 60/081203
;  PRIOR FILING DATE: 1998-04-09
;  PRIOR APPLICATION NUMBER: 60/081229
;  PRIOR FILING DATE: 1998-04-09
;  PRIOR APPLICATION NUMBER: 60/081955
;  PRIOR FILING DATE: 1998-04-15
;  PRIOR APPLICATION NUMBER: 60/081817
;  PRIOR FILING DATE: 1998-04-15
;  PRIOR APPLICATION NUMBER: 60/081819
;  PRIOR FILING DATE: 1998-04-15
;  PRIOR APPLICATION NUMBER: 60/081952
;  PRIOR FILING DATE: 1998-04-15
;  PRIOR APPLICATION NUMBER: 60/081838
;  PRIOR FILING DATE: 1998-04-15
;  PRIOR APPLICATION NUMBER: 60/082568
;  PRIOR FILING DATE: 1998-04-21
;  PRIOR APPLICATION NUMBER: 60/082569
;  PRIOR FILING DATE: 1998-04-21
;  PRIOR APPLICATION NUMBER: 60/082704
;  PRIOR FILING DATE: 1998-04-22
;  PRIOR APPLICATION NUMBER: 60/082804
;  PRIOR FILING DATE: 1998-04-22
;  PRIOR APPLICATION NUMBER: 60/082700
;  PRIOR FILING DATE: 1998-04-22
;  PRIOR APPLICATION NUMBER: 60/082797
;  PRIOR FILING DATE: 1998-04-22
;  PRIOR APPLICATION NUMBER: 60/082796
;     PRIOR FILING DATE: 1998-04-23
;     PRIOR APPLICATION NUMBER: 60/083336
;     PRIOR FILING DATE: 1998-04-27
;     PRIOR APPLICATION NUMBER: 60/083322
;     PRIOR FILING DATE: 1998-04-28
;     PRIOR APPLICATION NUMBER: 60/083392
;     PRIOR FILING DATE: 1998-04-29
;     PRIOR APPLICATION NUMBER: 60/083495
;     PRIOR FILING DATE: 1998-04-29
;     PRIOR APPLICATION NUMBER: 60/083496
;     PRIOR FILING DATE: 1998-04-29
;     PRIOR APPLICATION NUMBER: 60/083499
;     PRIOR FILING DATE: 1998-04-29
;     PRIOR APPLICATION NUMBER: 60/083545
;     PRIOR FILING DATE: 1998-04-29
;     PRIOR APPLICATION NUMBER: 60/083554
;     PRIOR FILING DATE: 1998-04-29
;     PRIOR APPLICATION NUMBER: 60/083558
;     PRIOR FILING DATE: 1998-04-29
;     PRIOR APPLICATION NUMBER: 60/083559
;     PRIOR FILING DATE: 1998-04-29
;     PRIOR APPLICATION NUMBER: 60/083500
;     PRIOR FILING DATE: 1998-04-29
;     PRIOR APPLICATION NUMBER: 60/083742
;     PRIOR FILING DATE: 1998-04-30
;     PRIOR APPLICATION NUMBER: 60/084366
;     PRIOR FILING DATE: 1998-05-05
;     PRIOR APPLICATION NUMBER: 60/084414
;     PRIOR FILING DATE: 1998-05-06
;     PRIOR APPLICATION NUMBER: 60/084441
;     PRIOR FILING DATE: 1998-05-06
;     PRIOR APPLICATION NUMBER: 60/084637
;     PRIOR FILING DATE: 1998-05-07
;     PRIOR APPLICATION NUMBER: 60/084639
;     PRIOR FILING DATE: 1998-05-07
;     PRIOR APPLICATION NUMBER: 60/084640
;     PRIOR FILING DATE: 1998-05-07
;     PRIOR APPLICATION NUMBER: 60/084598
;     PRIOR FILING DATE: 1998-05-07
;     PRIOR APPLICATION NUMBER: 60/084600
;     PRIOR FILING DATE: 1998-05-07
;     PRIOR APPLICATION NUMBER: 60/084627
;     PRIOR FILING DATE: 1998-05-07
;     PRIOR APPLICATION NUMBER: 60/084643
;     PRIOR FILING DATE: 1998-05-07
;     PRIOR APPLICATION NUMBER: 60/085339
;     PRIOR FILING DATE: 1998-05-13
;     PRIOR APPLICATION NUMBER: 60/085338
;     PRIOR FILING DATE: 1998-05-13
;     PRIOR APPLICATION NUMBER: 60/085323
;     PRIOR FILING DATE: 1998-05-13
;     PRIOR APPLICATION NUMBER: 60/085582
;     PRIOR FILING DATE: 1998-05-15
;     PRIOR APPLICATION NUMBER: 60/085700
;     PRIOR FILING DATE: 1998-05-15
;     PRIOR APPLICATION NUMBER: 60/085689
;     PRIOR FILING DATE: 1998-05-15
;     PRIOR APPLICATION NUMBER: 60/085579
;     PRIOR FILING DATE: 1998-05-15
;     PRIOR APPLICATION NUMBER: 60/085580
;     PRIOR FILING DATE: 1998-05-15
;     PRIOR APPLICATION NUMBER: 60/085573
;     PRIOR FILING DATE: 1998-05-15
;     PRIOR APPLICATION NUMBER: 60/085704
;     PRIOR FILING DATE: 1998-05-15
;     PRIOR APPLICATION NUMBER: 60/085697

  Query Match             22.7%;  Score 2058;  DB 6;  Length 525;
  Best Local Similarity   68.7%;  Pred. No. 1.2e-109;
  Matches  346;  Conservative   71;  Mismatches   87;  Indels    0;  Gaps    0;

Qy         64 ISRNIRSDFDKDGNWEAWGAWSQCSRTCGGGASYSLRRCLTGRNCEGQNIRYKTCSNEDC 123
              :||   ||:||:||  ||:|:  ||:|||||||||||||||:  :|||:|||:|||  ||
Db         22 SSRIARSEEDRDGLWDAWGPWSECSRTCGGGASYSLRRCLSSKSCEGRNIRYRTCSNVDC 81

Qy        124 PPDAEDFRAQQCSAYNDVQYQGHYYEWLPRYNDPAAPCALKCHAQGQNLVVELAPKVLDG 183
              ||:| ||||||||||:|||:: |  :|||||    |||    ||:||| |:|    ||||||||||||
Db         82 PPEAGDFRAQQCSAHNDVKHHGQFYEWLPVSNDPDNPCSLKCQAKGTTLVVELAPKVLDG 141

Qy        184 TRCNTDSLDMCISGICQAVGCDRQLGSNAKEDNCGVCAGDGSTCRLVRGQSKSHVSPEKR 243
              |||  |:|||||||||:||  |||  ||||    ||||||||  |||||  ||||||  |    :|    |
Db        142 TRCYTESLDMCISGLCQIVGCDHQLGSIVKEDNCGVCNGDGSTCRLVRGQYKSQLSAIKS 201

Qy        244 EENVIAVPLGSRSVRITVKGPAHLFIESKTLQGSKGEHSFNSPGVFVVENTTVEFQRGSE 303
              :: |:|:| |||  :|:  :|| ||::|:|||||:||:|    |  ||:|:|::|:|||:    :
Db        202 DDTVVALPYGSRHIRLVLKGPDHLYLETKTLQGTKGENSLSSTGTFLVDNSSVDFQKFPD 261

Qy        304 RQTFKIPGPLMADFIFKTRYTAAKDSVVQFFFYQPISHQWRQTDFFPCTVTCGGGYQLNS 363
              ::    :: |||  ||||  |  :  :  ||  :||  ||  ||    :|:|:||  ||:  |:||||||  |
Db        262 KEILRMAGPLTADFIVKIRNSGSADSTVQFIFYQPIIHRWRETDFFPCSATCGGGYQLTS 321

Qy        364 AECVDIRLKRVVPDHYCHYYPENVKPKPKLKECSMDPCPSSDGFKEIMPYDHFQPLPRWE 423
              |||  |:|    |||  |  ||||:|:|:||||:||:::|||:|:|:|:|:||||  :  ||||||
Db        322 AECYDLRSNRVVADQYCHYYPENIKPKPKLQECNLDPCPASDGYKQIMPYDLYHPLPRWE 381

Qy        424 HNPWTACSVSCGGGIQRRSFVCVEESMHGEIIQVEEWKCMYAPKPKVMQTCNLFDCPKWI 483
                 |||||: ||:||||  |:    |:||  :  :  :    ||||||||  ||    :  |  ||:|||||||:
Db        382 ATPWTACSSSCGGGIQSRAVSCVEEDIQGHVTSVEEWKCMYTPKMPIAQPCNIFDCPKWL 441

Qy        484 AMEWSQCTVTCGRGLRYRVVLCINHRGEHVGGCNPQLKLHIKEECVIPIPCYKPKEKSPV 543
              | |||  ||||||:||||||||||:|||  |  |||:|:  |  ||||||::|  ||||||||  ||
Db        442 AQEWSPCTVTCGQGLRYRVVLCIDHRGMHTGGCSPKTKPHIKEECIVPTPCYKPKEKLPV 501

Qy        544 EAKLPWLKQAQELEETRIATEEPT 567
              |||||| ||||||||     :|||:
Db        502 EAKLPWFKQAQELEEGAAVSEEPS 525


RESULT 14
US-09-978-192A-301
; Sequence 301, Application US/09978192A
; Patent No. US20020177553A1
; GENERAL INFORMATION:
;  APPLICANT: Ashkenazi, Avi
;  APPLICANT:  Baker, Kevin P.
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan
;  APPLICANT:  Ferrara, Napoleon
;  APPLICANT:  Filvaroff, Ellen
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Gerber, Hanspeter
;  APPLICANT:  Gerritsen, Mary E.
;  APPLICANT:  Goddard, Audrey
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, J. Christopher
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth J
;  APPLICANT:  Kljavin, Ivar J.
;  APPLICANT:  Kuo, Sophia S.
;  APPLICANT:  Napier, Mary A.
;  APPLICANT:  Pan, James;
;  APPLICANT:  Paoni, Nicholas F.
;  APPLICANT:  Roy, Margaret Ann
;  APPLICANT:  Shelton, David L.
;  APPLICANT:  Stewart, Timothy A.
;  APPLICANT:  Tumas, Daniel
;  APPLICANT:  Williams, P. Mickey
;  APPLICANT:  Wood, William I.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: P2630P1C9
;  CURRENT APPLICATION NUMBER: US/09/978,192A
;  CURRENT FILING DATE:  2001-10-15
;  PRIOR APPLICATION NUMBER: 09/918585
;  PRIOR FILING DATE: 2001-07-30
;  PRIOR APPLICATION NUMBER: 60/062250
;  PRIOR FILING DATE: 1997-10-17
;  PRIOR APPLICATION NUMBER: 60/064249
;  PRIOR FILING DATE: 1997-11-03
;  PRIOR APPLICATION NUMBER: 60/065311
;  PRIOR FILING DATE: 1997-11-13
;  PRIOR APPLICATION NUMBER: 60/066364
;  PRIOR FILING DATE: 1997-11-21
;  PRIOR APPLICATION NUMBER: 60/077450
;  PRIOR FILING DATE: 1998-03-10
;  PRIOR APPLICATION NUMBER: 60/077632
;  PRIOR FILING DATE: 1998-03-11
;  PRIOR APPLICATION NUMBER: 60/077641
;  PRIOR FILING DATE: 1998-03-11
;  PRIOR APPLICATION NUMBER: 60/077649
;  PRIOR FILING DATE: 1998-03-11
;  PRIOR APPLICATION NUMBER: 60/077791
;  PRIOR FILING DATE: 1998-03-12
;  PRIOR APPLICATION NUMBER: 60/078004
;  PRIOR FILING DATE: 1998-03-13
;  PRIOR APPLICATION NUMBER: 60/078886
;  PRIOR FILING DATE: 1998-03-20
;  PRIOR APPLICATION NUMBER: 60/078936
;  PRIOR FILING DATE: 1998-03-20
;  PRIOR APPLICATION NUMBER: 60/078910
;  PRIOR FILING DATE: 1998-03-20
;  PRIOR APPLICATION NUMBER: 60/078939
;  PRIOR FILING DATE: 1998-03-20
;  PRIOR APPLICATION NUMBER: 60/079294
;  PRIOR FILING DATE: 1998-03-25
;  PRIOR APPLICATION NUMBER: 60/079656
;  PRIOR FILING DATE: 1998-03-26
;  PRIOR APPLICATION NUMBER: 60/079664
;  PRIOR FILING DATE: 1998-03-27
;  PRIOR APPLICATION NUMBER: 60/079689
;  PRIOR FILING DATE: 1998-03-27
;  PRIOR APPLICATION NUMBER: 60/079663
;  PRIOR FILING DATE: 1998-03-27
;  PRIOR APPLICATION NUMBER: 60/079728
;  PRIOR FILING DATE: 1998-03-27
;  PRIOR APPLICATION NUMBER: 60/079786
;  PRIOR FILING DATE: 1998-03-27
;  PRIOR APPLICATION NUMBER: 60/079920
;  PRIOR FILING DATE: 1998-03-30
;  PRIOR APPLICATION NUMBER: 60/079923
;  PRIOR FILING DATE: 1998-03-30
;  PRIOR APPLICATION NUMBER: 60/080105
;  PRIOR FILING DATE: 1998-03-31
;  PRIOR APPLICATION NUMBER: 60/080107
;  PRIOR FILING DATE: 1998-03-31
;  PRIOR APPLICATION NUMBER: 60/080165
;  PRIOR FILING DATE: 1998-03-31
;  PRIOR APPLICATION NUMBER: 60/080194
;  PRIOR FILING DATE: 1998-03-31
;  PRIOR APPLICATION NUMBER: 60/080327
;  PRIOR FILING DATE: 1998-04-01
;  PRIOR APPLICATION NUMBER: 60/080328
;  PRIOR FILING DATE: 1998-04-01
;  PRIOR APPLICATION NUMBER: 60/080333
;  PRIOR FILING DATE: 1998-04-01
;  PRIOR APPLICATION NUMBER: 60/080334
;  PRIOR FILING DATE: 1998-04-01
;  PRIOR APPLICATION NUMBER: 60/081070
;  PRIOR FILING DATE: 1998-04-08
;  PRIOR APPLICATION NUMBER: 60/081049
;  PRIOR FILING DATE: 1998-04-08
;  PRIOR APPLICATION NUMBER: 60/081071
;  PRIOR FILING DATE: 1998-04-09
;  PRIOR APPLICATION NUMBER: 60/081195
;  PRIOR FILING DATE: 1998-04-08
;  PRIOR APPLICATION NUMBER: 60/081203
;  PRIOR FILING DATE: 1998-04-09
```

```
;  PRIOR APPLICATION NUMBER: 60/081229
;  PRIOR FILING DATE: 1998-04-09
;  PRIOR APPLICATION NUMBER: 60/081955
;  PRIOR FILING DATE: 1998-04-15
;  PRIOR APPLICATION NUMBER: 60/081817
;  PRIOR FILING DATE: 1998-04-15
;  PRIOR APPLICATION NUMBER: 60/081819
;  PRIOR FILING DATE: 1998-04-15
;  PRIOR APPLICATION NUMBER: 60/081952
;  PRIOR FILING DATE: 1998-04-15
;  PRIOR APPLICATION NUMBER: 60/081838
;  PRIOR FILING DATE: 1998-04-15
;  PRIOR APPLICATION NUMBER: 60/082568
;  PRIOR FILING DATE: 1998-04-21
;  PRIOR APPLICATION NUMBER: 60/082569
;  PRIOR FILING DATE: 1998-04-21
;  PRIOR APPLICATION NUMBER: 60/082704
;  PRIOR FILING DATE: 1998-04-22
;  PRIOR APPLICATION NUMBER: 60/082804
;  PRIOR FILING DATE: 1998-04-22
;  PRIOR APPLICATION NUMBER: 60/082700
;  PRIOR FILING DATE: 1998-04-22
;  PRIOR APPLICATION NUMBER: 60/082797
;  PRIOR FILING DATE: 1998-04-22
;  PRIOR APPLICATION NUMBER: 60/082796
;  PRIOR FILING DATE: 1998-04-23
;  PRIOR APPLICATION NUMBER: 60/083336
;  PRIOR FILING DATE: 1998-04-27
;  PRIOR APPLICATION NUMBER: 60/083322
;  PRIOR FILING DATE: 1998-04-28
;  PRIOR APPLICATION NUMBER: 60/083392
;  PRIOR FILING DATE: 1998-04-29
;  PRIOR APPLICATION NUMBER: 60/083495
;  PRIOR FILING DATE: 1998-04-29
;  PRIOR APPLICATION NUMBER: 60/083496
;  PRIOR FILING DATE: 1998-04-29
;  PRIOR APPLICATION NUMBER: 60/083499
;  PRIOR FILING DATE: 1998-04-29
;  PRIOR APPLICATION NUMBER: 60/083545
;  PRIOR FILING DATE: 1998-04-29
;  PRIOR APPLICATION NUMBER: 60/083554
;  PRIOR FILING DATE: 1998-04-29
;  PRIOR APPLICATION NUMBER: 60/083558
;  PRIOR FILING DATE: 1998-04-29
;  PRIOR APPLICATION NUMBER: 60/083559
;  PRIOR FILING DATE: 1998-04-29
;  PRIOR APPLICATION NUMBER: 60/083500
;  PRIOR FILING DATE: 1998-04-29
;  PRIOR APPLICATION NUMBER: 60/083742
;  PRIOR FILING DATE: 1998-04-30
;  PRIOR APPLICATION NUMBER: 60/084366
;  PRIOR FILING DATE: 1998-05-05
;  PRIOR APPLICATION NUMBER: 60/084414
;  PRIOR FILING DATE: 1998-05-06
;  PRIOR APPLICATION NUMBER: 60/084441
;  PRIOR FILING DATE: 1998-05-06
;  PRIOR APPLICATION NUMBER: 60/084637
;  PRIOR FILING DATE: 1998-05-07
;  PRIOR APPLICATION NUMBER: 60/084639
;  PRIOR FILING DATE: 1998-05-07
;  PRIOR APPLICATION NUMBER: 60/084640
;  PRIOR FILING DATE: 1998-05-07
;  PRIOR APPLICATION NUMBER: 60/084598
;  PRIOR FILING DATE: 1998-05-07
;  PRIOR APPLICATION NUMBER: 60/084600
;  PRIOR FILING DATE: 1998-05-07
;  PRIOR APPLICATION NUMBER: 60/084627
;  PRIOR FILING DATE: 1998-05-07
;  PRIOR APPLICATION NUMBER: 60/084643
;  PRIOR FILING DATE: 1998-05-07
;  PRIOR APPLICATION NUMBER: 60/085339
;  PRIOR FILING DATE: 1998-05-13
;  PRIOR APPLICATION NUMBER: 60/085338
;  PRIOR FILING DATE: 1998-05-13
;  PRIOR APPLICATION NUMBER: 60/085323
;  PRIOR FILING DATE: 1998-05-13
;  PRIOR APPLICATION NUMBER: 60/085582
;  PRIOR FILING DATE: 1998-05-15
;  PRIOR APPLICATION NUMBER: 60/085700
;  PRIOR FILING DATE: 1998-05-15
;  PRIOR APPLICATION NUMBER: 60/085689
;  PRIOR FILING DATE: 1998-05-15
;  PRIOR APPLICATION NUMBER: 60/085579
;  PRIOR FILING DATE: 1998-05-15
;  PRIOR APPLICATION NUMBER: 60/085580
;  PRIOR FILING DATE: 1998-05-15
;  PRIOR APPLICATION NUMBER: 60/085573
;  PRIOR FILING DATE: 1998-05-15
;  PRIOR APPLICATION NUMBER: 60/085704
;  PRIOR FILING DATE: 1998-05-15
;  PRIOR APPLICATION NUMBER: 60/085697

  Query Match             22.0%;  Score 2058;  DB 9;  Length 525;
  Best Local Similarity   68.7%;  Pred. No. 1.2e-185;
  Matches  346;  Conservative   71;  Mismatches   87;  Indels    0;  Gaps    0;

Qy         64 ISRNIRSDEDKPCNWDAWGPWSPCSRTCGGGASYSLRRCLTGRNCEGQNIRYKPCSNHPC 123
              :||  ||:||:   |||||||| ||||||||||||||||| : ::|||:||||: |||  |
Db         22 SSRIARSEEDRDGLWDAWGPWSECSRTCGGGASYSLRRCLSSKSCEGRNIRYRTCSNVDC 81

Qy        124 PPDAEDFRAQQCSAYNDVQYQGHYYEWLPRYNDPAAPCALKCHAQGQNLVVELAPKVLDG 183
              ||:| |||| ||:| |:|      |||||  |       | | | |  ||||||||||||
Db         82 PPEAGDFRALQCTAHNEVRHHLLTYEWLPVSNLTLNSPCLECLAKGGGLVVELAPKVLDG 141

Qy        184 TRCNTDSLDMCISGICQAVGCDRQLGSNAKEDNCGVCAGDGSTCRLVRGQSKSHVSPEKR 243
              ||| |:||||||||:|| ||||:||||  |||||||| |||||||||||| ||  |  |
Db        142 TRCYTESLDMCISGLCQIVGCDHQLGSTVKEDNCGVCNGDGSTCRLVRGQYKSQLSAIKS 201

Qy        244 EENVIAVPLGSRSVRITVKGPAHLFIESKTLQGSKGEHSFNSPGVFVVENTIVEFQRGSE 303
              :: |:| | ||| :|:::||| ||::|:||||| ||| |: | | ::|:|  | :||  :
Db        202 DDTVVALPYGSRHIRLVLKGPDHLYLETKTLQGTKGENSLSSTGTFLVDNSSVDFQKFPD 261

Qy        304 RQTFKIPGPLMADFIFKIRYTAAKDSVVQFFFYQPISHQWRQTDFFPCTVTCGGGYQLNS 363
              ::     ||| |||| |||      | |||:|||||:|:||:||||||  |||||||| |
Db        262 KEILEMAGPLTADFIVKIRNSGSAPSTVQFIFYQPIIHRWRETDFFPCSATCGGGYQLTS 321

Qy        364 AECVDIRLKRVVPDHYCHYYPENVKPKPKLKECSMDPCPSSDGFKEIMPYDHFQPLPRWE 423
              ||| |:|  ||| | ||||||||:|||||| || :|||| ||| | |||||   ||||||
Db        322 AECYDLRSNRVVADQYCHYYPENIKPKPKLQECNLDPCPASDGYKQIMPYDLYHPLPRWE 381

Qy        424 HNPWTACSVSCGGGIQRRSFVCVEESMHGEILQVEEWKCMYAPKPKVMQTCNLFDCPKWI 483
                |||||| ||||||| |   ||||   |    |||||||| ||      ||:||||||:
Db        382 ATPWTACSSSCGGGIQSRAVSCVEEDIQGHVTSVEEWKCMYTPKMPIAQPCNIFDCPKWL 441

Qy        484 AMEWSQCTVTCGRGLRYRVVLCINHRGEHVGGCNPQLKLHIKEECVIPIPCYKPKEKSPV 543
              | ||| |||||| ||||||||||:||| | ||| |  | ||||||::| ||||||| ||
Db        442 AQEWSPCTVTCGQGLRYRVVLCIDHRGMHTGGCSPKTKPHIKEECIVPTPCYKPKEKLPV 501

Qy        544 EAKLPWLKQAQELEETRIATEEPT 567
              |||||| ||||||||     |||:
Db        502 EAKLPWFKQAQELEEGAAVSEEPS 525


RESULT 15
US-09-999-832A-301
; Sequence 301, Application US/09999832A
; Publication No. US20020142706A1
; GENERAL INFORMATION:
;  APPLICANT: Ashkenazi, Avi
;  APPLICANT:  Baker, Kevin P.
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan
;  APPLICANT:  Ferrara, Napoleone
;  APPLICANT:  Filvaroff, Ellen
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Gerber, Hanspeter
;  APPLICANT:  Gerritsen, Mary E.
;  APPLICANT:  Goddard, Audrey
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, J. Christopher
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth J.
;  APPLICANT:  Kljavin, Ivar J.
;  APPLICANT:  Kuo, Sophia S.
;  APPLICANT:  Napier, Mary A.
;  APPLICANT:  Pan, James;
;  APPLICANT:  Paoni, Nicholas F.
;  APPLICANT:  Roy, Margaret Ann
;  APPLICANT:  Shelton, David L.
;  APPLICANT:  Stewart, Timothy A.
;  APPLICANT:  Tumas, Daniel
;  APPLICANT:  Williams, P. Mickey
;  APPLICANT:  Wood, William I.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: P2630P1C63
;  CURRENT APPLICATION NUMBER: US/09/999,832A
;  CURRENT FILING DATE:  2001-10-24
;  PRIOR APPLICATION NUMBER: 09/918585
;  PRIOR FILING DATE: 2001-07-30
;  PRIOR APPLICATION NUMBER: 60/062250
;  PRIOR FILING DATE: 1997-10-17
;  PRIOR APPLICATION NUMBER: 60/064249
;  PRIOR FILING DATE: 1997-11-03
;  PRIOR APPLICATION NUMBER: 60/065311
;  PRIOR FILING DATE: 1997-11-13
;  PRIOR APPLICATION NUMBER: 60/066364
;  PRIOR FILING DATE: 1997-11-21
;  PRIOR APPLICATION NUMBER: 60/077450
;  PRIOR FILING DATE: 1998-03-10
;  PRIOR APPLICATION NUMBER: 60/077632
;  PRIOR FILING DATE: 1998-03-11
;  PRIOR APPLICATION NUMBER: 60/077641
;  PRIOR FILING DATE: 1998-03-11
;  PRIOR APPLICATION NUMBER: 60/077649
;  PRIOR FILING DATE: 1998-03-11
;  PRIOR APPLICATION NUMBER: 60/077791
;  PRIOR FILING DATE: 1998-03-12
;  PRIOR APPLICATION NUMBER: 60/078004
;  PRIOR FILING DATE: 1998-03-13
;  PRIOR APPLICATION NUMBER: 60/078886
;  PRIOR FILING DATE: 1998-03-20
;  PRIOR APPLICATION NUMBER: 60/078936
;  PRIOR FILING DATE: 1998-03-20
;  PRIOR APPLICATION NUMBER: 60/078910
;  PRIOR FILING DATE: 1998-03-20
;  PRIOR APPLICATION NUMBER: 60/078939
;  PRIOR FILING DATE: 1998-03-20
;  PRIOR APPLICATION NUMBER: 60/079294
;  PRIOR FILING DATE: 1998-03-25
;  PRIOR APPLICATION NUMBER: 60/079656
;  PRIOR FILING DATE: 1998-03-26
;  PRIOR APPLICATION NUMBER: 60/079664
;  PRIOR FILING DATE: 1998-03-27
;  PRIOR APPLICATION NUMBER: 60/079689
;  PRIOR FILING DATE: 1998-03-27
;  PRIOR APPLICATION NUMBER: 60/079663
;  PRIOR FILING DATE: 1998-03-27
;  PRIOR APPLICATION NUMBER: 60/079728
;  PRIOR FILING DATE: 1998-03-27
;  PRIOR APPLICATION NUMBER: 60/079786
;  PRIOR FILING DATE: 1998-03-27
;  PRIOR APPLICATION NUMBER: 60/079920
;  PRIOR FILING DATE: 1998-03-30
;  PRIOR APPLICATION NUMBER: 60/079923
;  PRIOR FILING DATE: 1998-03-30
;  PRIOR APPLICATION NUMBER: 60/080105
;  PRIOR FILING DATE 1998-03-31
;  PRIOR APPLICATION NUMBER: 60/080107
;  PRIOR FILING DATE: 1998-03-31
;  PRIOR APPLICATION NUMBER: 60/080165
;  PRIOR FILING DATE: 1998-03-31
;  PRIOR APPLICATION NUMBER: 60/080194
;  PRIOR FILING DATE: 1998-03-31
;  PRIOR APPLICATION NUMBER: 60/080327
;  PRIOR FILING DATE: 1998-04-01
;  PRIOR APPLICATION NUMBER: 60/080328
;  PRIOR FILING DATE: 1998-04-01
;  PRIOR APPLICATION NUMBER: 60/080333
;  PRIOR FILING DATE: 1998-04-01
;  PRIOR APPLICATION NUMBER: 60/080334
;  PRIOR FILING DATE: 1998-04-01
;  PRIOR APPLICATION NUMBER: 60/081070
;  PRIOR FILING DATE: 1998-04-08
;  PRIOR APPLICATION NUMBER: 60/081049
;  PRIOR FILING DATE: 1998-04-08
;  PRIOR APPLICATION NUMBER: 60/081071
;  PRIOR FILING DATE: 1998-04-08
;  PRIOR APPLICATION NUMBER: 60/081195
;  PRIOR FILING DATE: 1998-04-08
;  PRIOR APPLICATION NUMBER: 60/081203
;  PRIOR FILING DATE: 1998-04-09
;  PRIOR APPLICATION NUMBER: 60/081229
;  PRIOR FILING DATE: 1998-04-09
;  PRIOR APPLICATION NUMBER: 60/081955
;  PRIOR FILING DATE: 1998-04-15
;  PRIOR APPLICATION NUMBER: 60/081817
;  PRIOR FILING DATE: 1998-04-15
;  PRIOR APPLICATION NUMBER: 60/081819
;  PRIOR FILING DATE: 1998-04-15
;  PRIOR APPLICATION NUMBER: 60/081952
;  PRIOR FILING DATE: 1998-04-15
;  PRIOR APPLICATION NUMBER: 60/081838
;  PRIOR FILING DATE: 1998-04-15
;  PRIOR APPLICATION NUMBER: 60/082568
;  PRIOR FILING DATE: 1998-04-21
;  PRIOR APPLICATION NUMBER: 60/082569
;  PRIOR FILING DATE: 1998-04-21
;  PRIOR APPLICATION NUMBER: 60/082704
;  PRIOR FILING DATE: 1998-04-22
;  PRIOR APPLICATION NUMBER: 60/082804
;  PRIOR FILING DATE: 1998-04-22
;  PRIOR APPLICATION NUMBER: 60/082700
;  PRIOR FILING DATE: 1998-04-22
;  PRIOR APPLICATION NUMBER: 60/082797
;  PRIOR FILING DATE: 1998-04-22
;  PRIOR APPLICATION NUMBER: 60/082796
;  PRIOR FILING DATE: 1998-04-23
;  PRIOR APPLICATION NUMBER: 60/083336
;  PRIOR FILING DATE: 1998-04-27
;  PRIOR APPLICATION NUMBER: 60/083322
;  PRIOR FILING DATE: 1998-04-28
;  PRIOR APPLICATION NUMBER: 60/083392
;  PRIOR FILING DATE: 1998-04-29
;  PRIOR APPLICATION NUMBER: 60/083495
;  PRIOR FILING DATE: 1998-04-29
;  PRIOR APPLICATION NUMBER: 60/083496
;  PRIOR FILING DATE: 1998-04-29
;  PRIOR APPLICATION NUMBER: 60/083499
;  PRIOR FILING DATE: 1998-04-29
;  PRIOR APPLICATION NUMBER: 60/083545
;  PRIOR FILING DATE: 1998-04-29
;  PRIOR APPLICATION NUMBER: 60/083554
;  PRIOR FILING DATE: 1998-04-29
;  PRIOR APPLICATION NUMBER: 60/083558
;  PRIOR FILING DATE: 1998-04-29
;  PRIOR APPLICATION NUMBER: 60/083559
;  PRIOR FILING DATE: 1998-04-29
;  PRIOR APPLICATION NUMBER: 60/083500
;  PRIOR FILING DATE: 1998-04-29
;  PRIOR APPLICATION NUMBER: 60/083742
;  PRIOR FILING DATE: 1998-04-30
;  PRIOR APPLICATION NUMBER: 60/084366
;  PRIOR FILING DATE: 1998-05-05
;  PRIOR APPLICATION NUMBER: 60/084414
;  PRIOR FILING DATE: 1998-05-06
;  PRIOR APPLICATION NUMBER: 60/084441
;  PRIOR FILING DATE: 1998-05-06
;  PRIOR APPLICATION NUMBER: 60/084637
;  PRIOR FILING DATE: 1998-05-07
;  PRIOR APPLICATION NUMBER: 60/084639
;  PRIOR FILING DATE: 1998-05-07
;  PRIOR APPLICATION NUMBER: 60/084640
;  PRIOR FILING DATE: 1998-05-07
;  PRIOR APPLICATION NUMBER: 60/084598
;  PRIOR FILING DATE: 1998-05-07
;  PRIOR APPLICATION NUMBER: 60/084600
;  PRIOR FILING DATE: 1998-05-07
;  PRIOR APPLICATION NUMBER: 60/084627
;  PRIOR FILING DATE: 1998-05-07
;  PRIOR APPLICATION NUMBER: 60/084643
;  PRIOR FILING DATE: 1998-05-07
;  PRIOR APPLICATION NUMBER: 60/085339
;  PRIOR FILING DATE: 1998-05-13
;  PRIOR APPLICATION NUMBER: 60/085338
;  PRIOR FILING DATE: 1998-05-13
;  PRIOR APPLICATION NUMBER: 60/085323
;  PRIOR FILING DATE: 1998-05-13
;  PRIOR APPLICATION NUMBER: 60/085582
;  PRIOR FILING DATE: 1998-05-15
;  PRIOR APPLICATION NUMBER: 60/085700
;  PRIOR FILING DATE: 1998-05-15
;  PRIOR APPLICATION NUMBER: 60/085689
;  PRIOR FILING DATE: 1998-05-15
;  PRIOR APPLICATION NUMBER: 60/085579
;  PRIOR FILING DATE: 1998-05-15
;  PRIOR APPLICATION NUMBER: 60/085580
;  PRIOR FILING DATE: 1998-05-15
;  PRIOR APPLICATION NUMBER: 60/085573
;  PRIOR FILING DATE: 1998-05-15
;  PRIOR APPLICATION NUMBER: 60/085704
;  PRIOR FILING DATE: 1998-05-15
;  PRIOR APPLICATION NUMBER: 60/085697

  Query Match             22.0%;  Score 2058;  DB 9;  Length 525;
  Best Local Similarity   68.7%;  Pred. No. 1.2e-109;
  Matches  346;  Conservative   71;  Mismatches   87;  Indels    0;  Gaps    0;

Qy         64 TCRNTESDFFKFGNWDAWCFWSGCSKTCGGGASYSLRRCLTGRNCEGQNIRYKTCSNHDC 123
              :||  ||:||:|| ||||| ||:|||||||||||||||||||: ::|||:||||:|||| ||
Db         22 SSKTARSEEDKEGLWDAWGPWSECSRTCGGGASYSLRRCLSSKSCEGRNIRYRTCSNVDC 81

Qy        124 PPDAEDFRAQQCSAYNDVQYQGHYYEWLPRYNDPAAPCALKCHAQGQNLVVELAPKVLDG 183
              ||:| ||||||||||:|||:: | :|||||   |||   ||:||| |:|   |||||||||||
Db         82 PPEAGDFRAQQCSAHNDVKHHGQFYEWLPVSNDPDNPCSLKCQAKGTTLVVELAPKVLDG 141

Qy        184 TRCNTDSLDMCISGICQAVGCDRQLGSNAKEDNCGVCAGDGSTCRLVRGQSKSHVSPEKR 243
              ||| |:|||||||||:|| |||| ||||   ||||||||  |||||||||||| || :|   |
Db        142 TRCYTESLDMCISGLCQIVGCDHQLGSTVKEDNCGVCNGDGSTCRLVRGQYKSQLSAFKS 201

Qy        244 EENVIAVPLGSRSVRITVKGPAHLFIESKTLQGSKGEHSFNSPGVFVVENTTVEFQRGSE 303
              :: |:|:| ||| :|: :||| ||::|:|||||:|||:| :| | |:|:|::|:||: :
Db        202 DDTVVALPYGSRHIRLVLKGPDHLYLETKTLQGTKGENSLSSTGTFLVDNSSVDFQKFPD 261

Qy        304 RQTFKIPGPLMADFIFKTRYTAAKDSVVQFFFYQPISHQWRQTDFFPCTVTCGGGYQLNS 363
              :: :: ||| |||| | | : : || ||| ||||| |:||:|||||||: |||||||| |
Db        262 KEILRMAGPLTADFIVKIRNSGSADSTVQFIFYQPIIHRWRETDFFPCSATCGGGYQLTS 321

Qy        364 AECVDIRLKRVVPDHYCHYYPENVKPKPKLKECSMDPCPSSDGFKEIMPYDHFQPLPRWE 423
              ||| |:|  ||| | ||||||||:||||||:||::||||:|||:|:||||| : ||||||
Db        322 AECYDLRSNRVVADQYCHYYPENIKPKPKLQECNLDPCPASDGYKQIMPYDLYHPLPRWE 381

Qy        424 HNPWTACSVSCGGGIQRRSFVCVEESMHGEILQVEEWKCMYAPKPKVMQTCNLFDCPKWI 483
```

```
            ||:||  |||:||| ::  :|:| : | :  |:||||| ||  : | |.: |.:|.:
Db      382 ATPWIACSSSCGGGIQSRAVSCVEEDIQGHVISVEEWKCMYTPKMPIAQPCNIPPCPKWL 441

Qy      484 AMEWSQCTVISGRGLRYRVVLCINHRGLHVGGSNPQLKLHIKEECVIPIPCYKLKEKSFV 543
            | ||| .||:|:||||||| |: :||| | |:|:| | |||:|:::  |:|:|:|:| .
Db      442 AQPWSPCTVIGGQGCRYRVVLCTDHRGMHPGGCSPKTKPHIKEECIVPIPCYKPKEKLSV 501

Qy      544 EAKLPWLKQAQELEEIRIAIEEPL 567
            |:|||:| ||||||.||     : |::
Db      502 EAKLPWEKQAQELEEGAAVSEEPS 525

Search completed: March 22, 2003, 13:38:27
Job time : 37 secs
```